PÉTER ÁGOSTON

POINT DEFECT AND SURFACE PROPERTIES OF In_2O_3 AND SnO_2: A COMPARATIVE STUDY BY FIRST-PRINCIPLES METHODS

Zur Erlangung des akademischen Grades des Doktors der Ingenieurwissenschaften
(Dr.-Ing.) genehmigte Dissertation
vorgelegt von Dipl.-Ing. Péter Ágoston
Geburtsort: Tirgu Mures, Rumänien

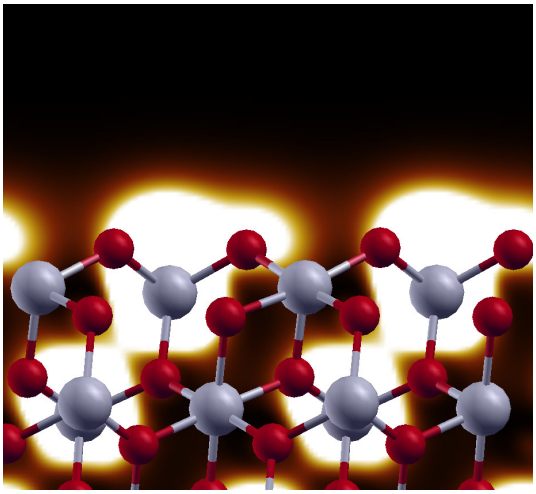
Fachgebiet Materialmodellierung
Fachbereich Material- und Geowissenschaften
Technische Universität Darmstadt

Referent: Prof. Karsten Albe
Korreferent: Prof. Heinz von Seggern

Tag der Einreichung: 19. Mai 2011
Tag der Mündlichen Prüfung: 6. Juli 2011
Erscheinungsort: Darmstadt
Erscheinungsjahr: 2011
D 17

POINT DEFECT AND SURFACE PROPERTIES OF In_2O_3 AND SnO_2: A COMPARATIVE STUDY BY FIRST-PRINCIPLES METHODS

PÉTER ÁGOSTON



Dissertation

May 2011

On the cover: Empty states STM simulation of a stoichiometric SnO_2-101 surface reconstruction.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

χ_A	electron affinity
$O_{i,db}$	oxygen dumbbell interstitial
O_i	oxygen interstitial
V_{In}	indium vacancy
V_O	oxygen vacancy
V_{Sn}	tin vacancy
E_f	Fermi energy
I_p	ionization potential
a-IZO	amorphous indium zinc oxide
ATO	antimony doped tin oxide
AZO	aluminum doped zinc oxide
BM shift	Burstein-Moss shift
CB	conduction band
CBM	conduction band minimum
DOS	electronic density of states
FCM	force constant matrix
FHA	full harmonic approximation
FTO	fluorine doped tin oxide
GIZO	gallium indium zinc oxide
ITO	tin doped indium oxide
IZO	indium zinc oxide
LEIS	low energy ion scatterin
nc-AFM	non contact atomic force microscopy

pDOS	phonon density of states
STM	scanning tunneling microscopy
TC	transparent conductor
TCO	transparent conducting oxide
TH	Tersoff-Hamann
TM	transition metal
UPS	ultra violet photoelectron spectroscopy
VB	valence band
VBM	valence band minimum
XPS	X-ray photoelectron spectroscopy

Part I

INTRODUCTION

1

MOTIVATION

Transparent conducting oxides (TCO) are used as electrodes for displays, solar cells and organic electronics, as gas sensors, catalyst, heat shields or for DNA detection. The increasing demand for environmentally friendly power production and the prospect of fully transparent electronics have additionally driven the research on TCO materials in recent years. Many basic properties of TCO materials are still controversially discussed or simply unknown. In addition, new materials with enhanced transparency, higher conductivity, stability and/or new functionalities like *e.g.* magnetism are emerging. Furthermore, many researchers explore TCO materials on the nano-scale where a thorough knowledge of bulk and surface properties is indispensable.

This thesis sumarizes results obtained within subproject C2 (Atomistic computer simulations on point defects and their mobility in metal oxides) of SFB-595 of DFG (Fatigue of Functional Materials [1])

The aims of C2 are to elucidate potential degradation mechanisms at the interface of organic light emitting diodes with tin doped indium oxide (ITO) anodes. It was suggested that out-diffusion of oxygen or inhomogeneous electronic injection barriers of the TCO electrode could lead to a degradation of the organic material [1, 2]. In order to assess these degradation mechanisms a detailed knowledge of point defects, their kinetics and surface properties of the contact material are necessary. Thus, the following questions were addressed:

- **Defect equilibria -** Which types of point defects can be present in In_2O_3? (Chapter 5)
- **Doping limits -** What is the role of *n*-type doping and how does it change the defect equilibria within the material, *e.g.* in ITO? (Chapter 8)
- **Kinetics -** How fast is the kinetics of the point defects and is it sufficient to cause oxygen leakage from the anode at operating temperatures? (Chapter 9)
- **Surface stability of In_2O_3 -** Which are the stable surface structures and surface stoichiometries of In_2O_3? (Sections 12.3.1 and 12.3.2)
- **Surface stability of ITO -** What is the influence of *n*-type doping on the surface stability and stoichiometry? (Section 12.6)

- **Contact properties -** How does the the work function (injection barrier) vary with surface orientation, termination and composition? (Section 12.7.2)

In addition, the above mentioned aspects were also adressed for an alternative material, namely SnO_2 (Chapters 7 10 and 11).

The first part of the dissertation deals with the thermodynamic stability of point defects. This is done by first surveying possible defect geometries in order to identify relevant structures. The formation energies of selected defects are then further refined by estimating their formation entropies as well as by using more sophisticated DFT approximations like *hybrid-functionals*.

The second part deals with the kinetic properties of defects. The migration barriers of a range of defects is calculated in various charge states in order to estimate the diffusion constants and equilibration temperatures.

The third part finally turns to the surface properties In_2O_3 and SnO_2. First, the controversially discussed thermodynamic ground state properties of SnO_2 surfaces are resolved and complemented. Then a detailed survey on the stability and structure of surfaces is conducted for In_2O_3 surfaces. The influence of water, strain, dopants and band bending are discussed and linked to experimental observations.

Finally, the properties of bulk defects and surfaces are compared for the two materials. The results of this thesis provide a sound and comprehensive view on the microscopic processes in In_2O_3 and SnO_2. The information contained will be relevant in the fields of electrical properties, diffusion, contact properties, gas sensing and doping.

2

TRANSPARENT CONDUCTORS

2.1 TECHNOLOGIES AND APPLICATIONS

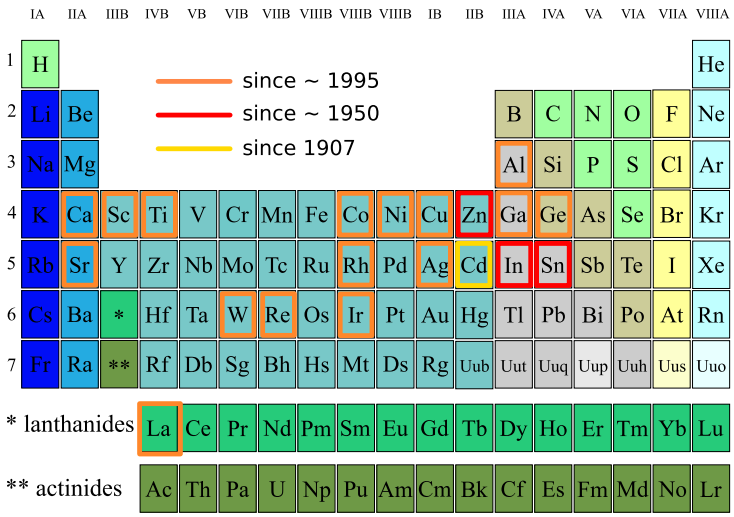


Figure 1: Periodic table of elements with boxes marking the relevant cations for transparent conducting oxide material. (Adopted from Ref. [3])

Bixbyite indium oxide (In_2O_3) and Cassiterite tin oxide (SnO_2) are members of the transparent conducting oxides (TCO) [4]. Their key properties are a nearly metal-like conductivity and good transparency for visible light. This combination of material properties is scarce in nature. Exemplified by the vast number of light reflecting metallic elements and alloys as well as the transparent insulators, the existence of one property usually excludes the existence of the other. Only a very limited number of compounds, most of which are oxides of the late transition and post transition metals, can exhibit conductivities up to $> 10^4$ Scm^{-1} while maintaining a transparency of $> 80\%$ in the visible range [5].

The lack of available materials systems has been stimulating the research since roughly the mid nineties of the last century when the number of studies of TCO systems has dramatically increased [3]. Fig. 1 shows the periodic table of elements with the typical TCO elements marked within. It is remarkable that TCOs have been around and also used since the year 1907. Especially, during World War II they secretly received considerable attention for the use in window defrosting and defogging devices which made high altitude bombing possible at that time. Nowadays most of the TCOs are produced for so called "green" technologies. The new impetus has been given by the growing concerns about changes of the world climate. Since then, research on TCOs has dramatically expanded because of its key role in power saving and "green" power producing applications.

2.1.1 *Light and heat control*

By quantity the most important application of TCOs are fluorine doped SnO_2 (FTO) glass coatings, which are usually applied by a chemical vapor deposition (CVD) process during the float-glass production [3, 6, 7]. These so-called low emittance ("low-e") glasses are transparent infrared reflectors [8] and allow solar control and energy saving heat management for buildings as well as for transport vehicles. Windows with tin oxide coatings are efficient in preventing radiative heat loss due to tin oxide's low emissivity of about 0.16 [9]. The plasma wavelength on the one side and the fundamental band-gap absorption of the TCO on the other side limit the transparent frequency range of the material, which can be tuned by selection of the TCO and its doping concentration. This enables the material to reflect IR-light in order to keep the buildings interior cool in summer. In winter it minimizes heat losses through windows as well (Fig. 2 [10]). In 2010 the world wide production of coated low-e glass has grown on the order of $10^7 m^2$ [11] mainly because of the rapid development in China [12], while in Europe or US the market for *low-e* glasses has already saturated. This basic technology has been established early and currently not a very active field of research. On the other side TCOs are an essential part of other and new functional window designs like *e.g.* electrochromic windows [13, 14, 15, 16]. Here, the idea is to change the absorption characteristics of a window within a few minutes or seconds by applying a voltage. In such devices it is necessary that the active layer is electrically contacted from both sides by materials with extraordinary good transparency. Finally, thermal and solar control is

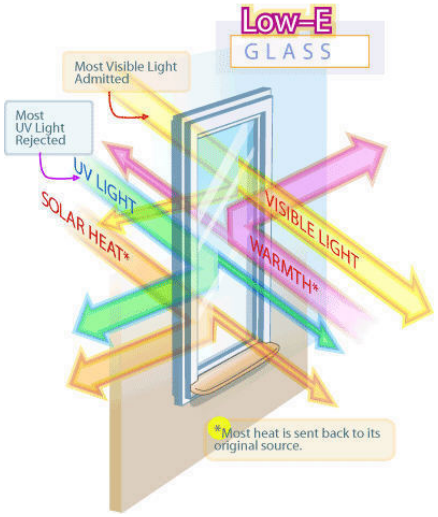


Figure 2: Illustration of a low-e glass

also desirable in the case of non-window facades of buildings. For such applications, colorless paints using TCO nanoparticles are currently developed [17]. These paints make rooftops reflecting for the IR part of the electromagnetic spectrum and avoid heating of the buildings interior.

2.1.2 *Electronic devices*

The second largest market for TCOs is provided by flat panel displays (FPD) [18]. Besides stationary displays, the increasing number of mobile and hand-held devices has lead to a growth of this branch [19]. Crystalline ITO is mainly used for these technologies as transparent electronic contact material due to its superior properties (transparency and electric conductivity) [20, 21, 22]. ITO is used as electrical contact material for individual pixels in the liquid crystal displays (LCD). A typical LCD architecture is illustrated in Fig.3 [23], where two ITO layers are employed. Currently, there is a large research focus on additional and/or alternative materials for replacing ITO in display technologies. This development is not only driven by the shortages of indium supply (see Sect. 2.2.7), but also by the changing technological demands in this field. Low temperature deposition methods are explored for new flexible electronics on plastic supports [24]. The need for low processing temperatures and good surface morphology as well as etching behavior or thermal stability has lead to the development of amorphous TCOs. Amorphous TCOs can be fabricated using conventional TCOs such as In_2O_3 [25, 26], but the physical properties and especially the thermal stability are, however, superior when amorphous indium zinc oxide (a-IZO) is used instead [27, 28, 29]. This new material is now becoming a competitor for ITO-based contact materials. Amorphous TCOs are showing in many aspects a superior behavior, and the effects of the broken translational symmetry have an unexpectedly low impact on the electron mobility. Homogeneity in combination with a considerable electron mobility is also of great interest for oxide thin film transistors (TFT), where again a-IZO is the most promising material [30].

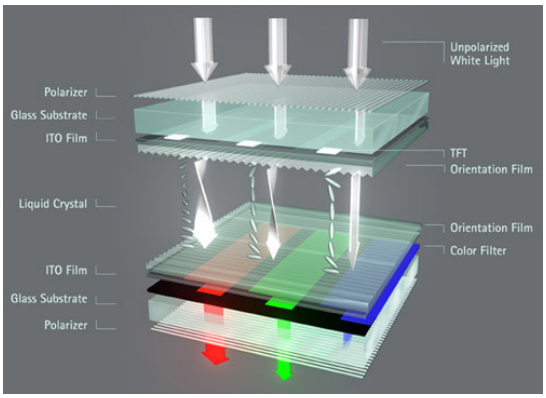


Figure 3: The architecture of a liquid crystal display

Finally, the holy grail of the display-related TCO technologies is the development of transparent *p*–*n* junctions. Only with this device the field of transparent electronics could be fully developed. While most of the TCOs are only *n*-type wide-gap semiconductor materials, a series of several different *p*-type materials has been pro-

posed, recently. Also transparent *p*–*n* heterojunctions have already been constructed, [31, 30, 22] which were, however, still far from being optimal.

Although the search for good *p*-type TCO materials presently receives most attention, further development in the field of *n*-type materials is indispensable. For present FPD applications ITO is still well suited. In the future, however, it will be necessary to have materials with enhanced transparency and conductivity for new high performance devices. This will only be possible by improving the electronic carrier mobility of materials with high carrier concentrations or vice versa. A first approach is to combine the traditional TCOs (In_2O_3, SnO_2, ZnO, CdO and Ga_2O_3) in ternary or quaternary compounds [32]. Promising materials are the ternary spinels, like *e.g.* $SnCd_2O_4$ [33, 34, 35]. This material is extremely heat resistant and has an excellent conductivity and transparency. It is used for the fabrication of high performance CdTe solar cells [36].

2.1.3 *Photovoltaics*

In comparison to window applications and FPDs, the TCO market for power production and, specifically, for photovoltaics (PV) devices is smaller. This segment is, however, the fastest growing field [3]. At present, 93% of the world wide produced PV elements are still based on silicon technology, where traditionally no TCOs were needed for contacts. There are two evolutions bringing TCOs into this field. On the one side, silicon based technologies nowadays apply TCOs in addition to silver strips as electrical contacts [3]. Fig. 4 shows two technologically significant silicon solar cell architectures (a. and b.), both of which make use of one or several TCO contact layers. Often, different TCO layer are used in combination, where *e.g.* one is highly conductive whereas the other is more insulating. TCOs have, therefore, become an important part also in most of the silicon based solar cell technologies [37]. A second development is the increasing interest in the so called *second-generation* or thin-film solar cell technology [38, 39]. Two of the most promising material systems are shown in Fig. 4 ((c): $Cu(InGa)Se_2$ (CIGS) and (d.): CdTe). For thin-film solar cell technologies, TCOs have been an essential part of the design since their beginnings. Several different TCOs are employed for this purpose. Typically, CIGS solar cells make use of a ZnO:Al contact layer, while FTO and more recently the multicomponent Cd_2SnO_4 are employed for CdTe technologies. These thin-film technologies have the potential of ultimately leading to low cost and high throughput fabrication [38, 39]. There are also several drawbacks using these technologies, like low abundance of the materials (Se,Te,In) or toxicity (Cd). While these issues are not settled, they will probably not be an inhibiting factor for thin-film technologies since new materials are steadily developed. In the future, it is likely that thin-film technologies will dominate the market over crystalline silicon solar cells. The world-wide leading manufacturer of solar-cells (First solar) has recently achieved an annual production output of devices which is on the order of Gigawatts [40], comparable to the power output of a nuclear

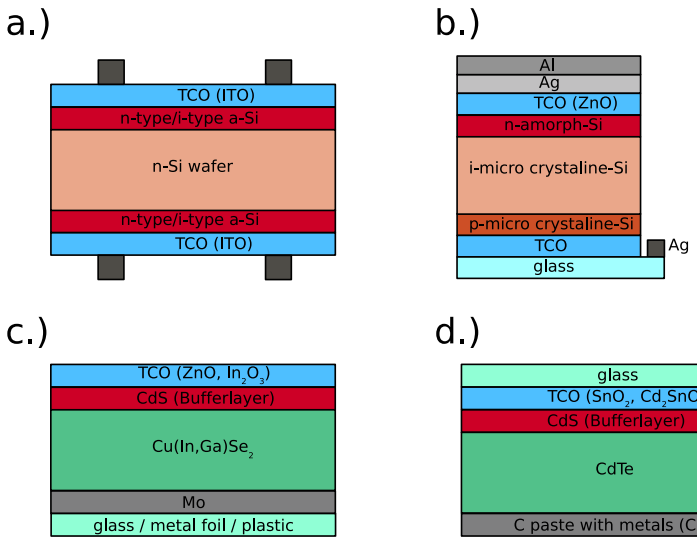


Figure 4: a.) Sanyo HIT-cell b.) amorphous-Si nip cell c.) CIGS cell d.) CdTe cell

reactor. Together with new production technologies, which involve roll-to-roll or printing techniques, cheaper solar-cells will finally become an economically competitive power source and provide a large and stable market for TCOs.

2.1.4 *Organic electronics*

An even cheaper and faster production of solar cells is expected for devices which use organics as absorber materials [41, 42, 43]. Also for these devices TCOs (mainly ITO) are heavily used. In this field several alternatives are tested due to the requirement of flexibility and good electronic injection. Amorphous TCOs like a-IZO are again good candidates for this purpose. At present, organic photovoltaic devices (OPV) still suffer from extremely low efficiencies, which are rising only slowly as a function of time compared with other technologies.

The inverse technology using organic materials as light emitting diodes (OLED) is far more advanced and relevant at this point [44]. The low power consumption makes this technology currently attractive as a future source of lighting. OLEDs can be deposited onto large flexible substrates and, consequently, emit light from the whole of the substrate's area and not from a restricted point or line source. Besides power saving aspects, this technology gives room for design aspects, too. Major challenges in the lighting sector are the lifetimes of such devices, which can be very good when small organic molecules are used as the light source and somewhat inferior in the case of polymeric materials. In general, however, there are still problems related to the lifetimes of the blue emitter materials [2].

2.1.5 *Gas sensors*

Typical TCO materials show pronounced compositional variations with changing environmental conditions at their surfaces but also in the bulk. This effect can be exploited for gas sensing applications [45, 46, 47]. In the case of SnO_2-based Taguchi Sensors [48, 49], the response is assumed to be only a surface effect in contrast to TiO_2 [50, 51]. Similarly, In_2O_3 can be used as gas sensor, especially for oxidizing gases such as ozone (O_3), NO_2 [52, 53, 54, 55] and Cl_2 [56], but also for reducing gases like CO [57]. A possible gas response mechanism is schematically depicted in Fig. 5. Negatively charged oxygen species ionosorb on the surface of the material at higher oxygen chemical potentials and cause a band-bending (depletion) in the boundary regions of the *n*-type material. This leads to a decreased electron concentration and increased barriers for electrons at the grain boundaries. As a result, the resistivity of the material increases and can indicate the presence of an oxidizing gas. This basic concept is used for the design of sensors, which are sensitive to oxidizing but also combustible and reducing gases. Such devices are important in the context of environmental and industrial control.

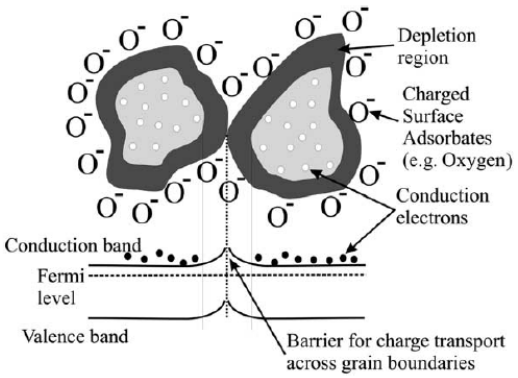


Figure 5: Schematic illustration of the surface gas response in SnO_2 based Taguchi sensors taken from Ref [51].

2.2 PHYSICS OF TRANSPARENT CONDUCTORS

In general, there are three conceptual routes to create a transparent conductor (TC). It is possible to take a well conducting and metallic material like *e.g.* silver and improve its transparency by reducing its layer thickness so that optical absorption is minimized, while the good conductivity is retained. This route is used *e.g.* in ZnO-Ag-ZnO layered systems or simple metal films of Au [58, 14, 6]. Problems with this technologies are, however, a bad uniformity of thin layers due to the wetting properties of metals. Furthermore, metal films are soft and easily oxidize so that often additional capsulation is necessary.

A second possibility is to use organic materials [59]. Problems are again related to the thermodynamic stability and reactivity of these materials. Another very promising and in many aspects ideal material is a single sheet of graphene [60], which can already be obtained for large areas.

The third and most successful method is to use an insulating material (*e.g.* metal oxide) and take measures to increase its electric conductivity. Typical insulating materials have large band-gaps (>2-3 eV) leading to a sufficiently good transparency in the visible range of the electromagnetic spectrum. The problem is, however, that even in a material with a band gap as low as $\sim$ 0.6 eV (Germanium) the electrical conductivity is unacceptably low for any electronic application due to low electronic carrier concentrations. In order to create the desired free carriers, doping has to be employed until the Fermi energy is moved to the valence band maximum (VBM) for a high concentration of hole carriers or to the conduction band minimum (CBM) for electron carriers. The connection between the electron n_e and hole n_h concentration as a function of the Fermi energy E_f is given by the product of the electronic density of states $D(E)$ and the Fermi occupation function $f(E, T, E_f)$ integrated over the relevant energy ranges for holes and electrons, respectively [61],

$$n_e = \int_{CBM}^{\infty} D(E) f(E, T, E_f) dE, \tag{2.1}$$

$$n_h = \int_{VBM}^{-\infty} D(E)[1 - f(E, T)] dE, \tag{2.2}$$

with

$$f(E, T) = \frac{1}{1 + e^{(E - E_f)/k_B T}}. \tag{2.3}$$

The conductivity of the material σ_{el} is then determined by the product of the majority carrier mobility and concentration,

$$\sigma_{el} = e(\mu_e n_e + \mu_h n_h). \tag{2.4}$$

2.2.1 *Wide gap materials*

A wide gap metal oxide which allows the doping with at least one carrier type can thus be converted into a transparent conducting oxide (TCO). The TCOs, which are considered in the present study, show a strong propensity towards *n*-type behavior while *p*-type doping yet remains a challenge. Alternative wide gap materials such as nitrides, selenides or sulfides can also be converted into conductive and transparent materials. Especially, gallium nitride (GaN) has attracted much attention [62, 63, 64] due to the possibility of both, *n*- and *p*-type doping. The superiority of oxides in comparison with nitrides is, however, the higher free electron concentration. Typical and most commonly used TCO materials are In_2O_3, SnO_2, ZnO, CdO and Ga_2O_3 (mostly amorph).

2.2.2 *Structural and thermodynamic properties of TCOs*

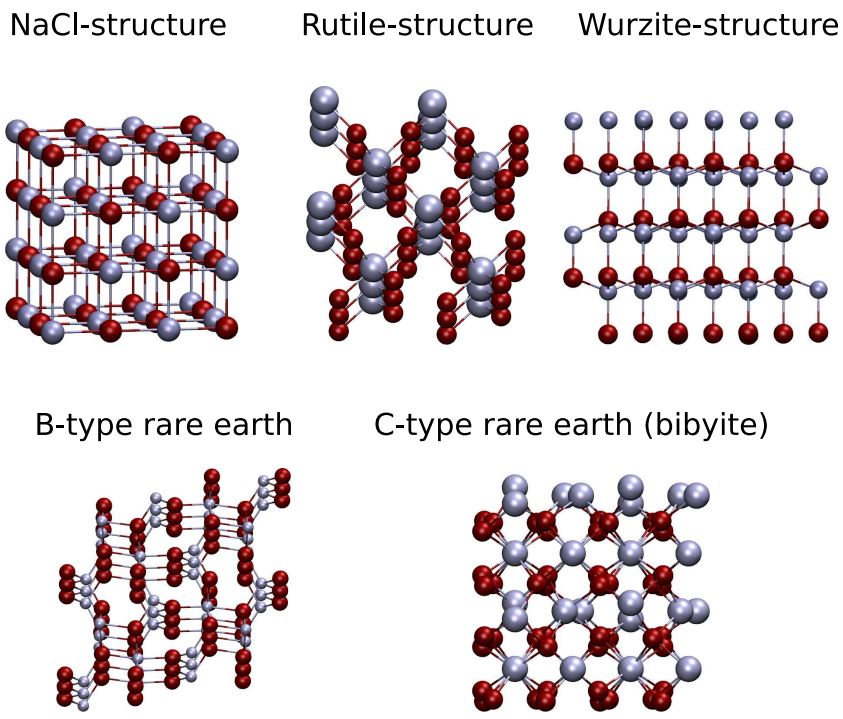


Figure 6: Crystal structures of the most common TCO materials.

Traditional TCO materials with late transition/post transition metal cations exhibit a variety of different crystal structures. Fig. 6 shows the prototype structures: rocksalt,

rutile, wurzite, B-type and C-type rare earth crystal structures. Table 1 lists structural properties and the heats of formation of ZnO, CdO, SnO_2, GeO_2, Ga_2O_3 and In_2O_3 along with the coordination of anions and cations. In general, there are compounds with tetrahedrally and octahedrally coordinated cations as well as tetrahedrally and trigonally coordinated anions. The most important structure with tetrahedrally coordinated cations is the ZnO wurzite lattice. Also, in the case of Ga_2O_3 a part of the cations have a close-to-tetrahedral coordination, whereas other cations have a more or less distorted octahedral coordination. Only in the case of CdO, a perfect octahedral coordination of cations is preserved, whereas it is distorted in the case of the rutile-SnO_2/GeO_2, β-Ga_2O_3 and bixbyite-In_2O_3 lattices. The oxygen anions only have octahedral coordination in CdO, whereas in wurzite, bixbyite and β-Ga_2O_3 there is an increasingly distorted tetrahedral coordination. In β-Ga_2O_3 there are three different oxygen sites having coordinations between tetrahedral and trigonal planar. In the case of the rutile lattice there is only a trigonal planar coordination for anions.

Table 1: Compilation of structural properties of different TCO materials. ΔH^f: compound heat of formation (eV/f.u) taken from Ref.[65]. $a, b, c, \alpha, \beta, \gamma$: lattice parameters, SG : space group number. Crystallographic data taken from Refs.[66, 65, 67, 68, 69] in the order the materials are listed.

TCO	ZnO	CdO	Ga_2O_3	In_2O_3	SnO_2
SG No.	186	225	12	206	136
symmetry	$P6_3mc$	$Fm\bar{3}m$	C2/m	$Ia\bar{3}$	$P4_2/mnm$
structure	wurzite	rocksalt	B-type	bixbiyte	rutile
cation coord.	tet.	oct.	tet./oct.	oct.	oct.
anion coord.	tet.	oct.	tet./tri.	tet.	tri.
a	3.25	4.70	12.23	10.117	4.74
b	3.25	4.70	3.04	10.117	4.74
c	5.21	4.70	5.80	10.117	3.19
α	90	90	90	90	90
β	90	90	103.7	90	90
γ	130	90	90	90	90
ΔH^f	-3.63	-2.68	-11.29	-9.59	-5.99

It is also noteworthy that some of the TCO materials exhibit several different polytypes. Zinc oxide can exist in the wurzite and zinc blende modification as well as in a rocksalt modification under high pressures [70], similar to CdO. The sesquioxides (In_2O_3 and Ga_2O_3) can exist in A-, B- and C-type rare earth modifications [71]. In the case of gallia all three modifications are possible while for In_2O_3, the monoclinic B-type modification has not been observed so far. In addition, the post transition metals Sn, Ge, In and Ga are able to form suboxides SnO ($\Delta H_f = -2.91$ eV), GeO ($\Delta H_f = -2.71$ eV) and Ga_2O ($\Delta H_f = -3.69$ eV) [65] which are, however, metastable *

*In the case of SnO it is not yet clear whether it forms part of the equilibrium phase diagram [72]

(In_2O is gas forming at high temperatures [71]). In contrast, the late transition metals Zn and Cd are able to form peroxide phases [73, 74] indicating their close relation to the alkaline earth metals. Table 1 lists only the ground state modifications of the respective materials.

2.2.3 *Electronic structure of TCOs*

The archetypical TCO materials are band insulators with moderate band gaps ($\sim$ 3 eV). The bonding is, however, predominantly ionic and the valence band (VB) and conduction band (CB) are, therefore, mainly anion and cation derived. The materials listed in the previous section exhibit only *n*-type behavior. The creation of electron holes in the VB of these materials seems to be intrinsically hindered by internal compensation processes and the Fermi energy appears to be confined within the upper part of the band gap under any condition. Especially in the case of ZnO [75, 76, 77, 78, 79], but also for SnO_2 [80] and In_2O_3 [81], the existence of *p*-type conductivity has been reported by several groups. These reports, however, tend to be not reproducible and in some cases they have been revealed to be artificial [82]. The reason for this is rooted in the electronic structure of these materials [83]. Fig. 7

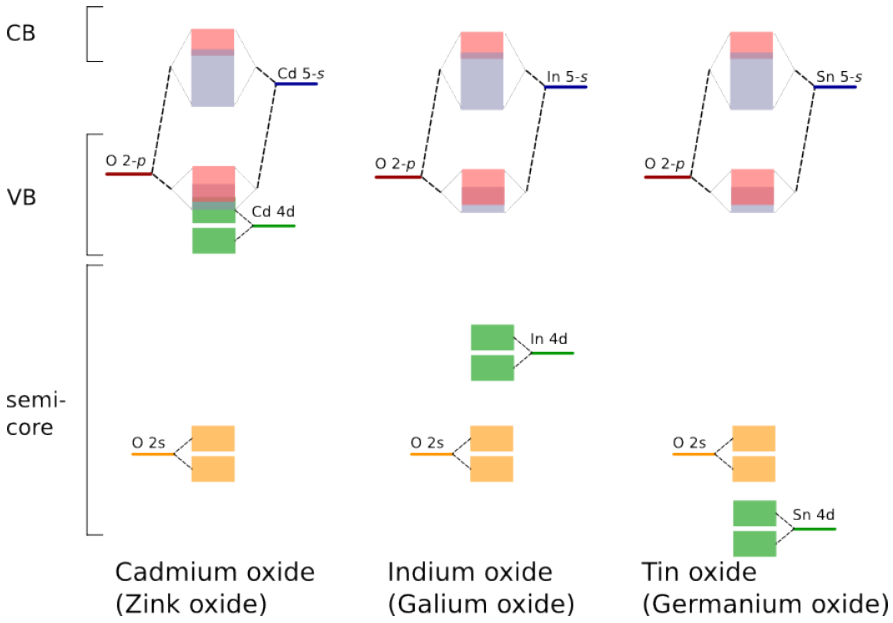


Figure 7: Schematic illustration of the TCOs electronic structures.

schematically illustrates the electronic structures of the oxides of the two late transition metal cations Zn and Cd as well as the four post transition metals Ge, Ga, In and Sn. The major difference between the electronic structure of these oxides is the location

of the occupied *d*-states which is at increasing binding energies when going to the right in the periodic table of elements. This effect is caused by the increasing core charge of the cations. In the case of ZnO and CdO the *d*-states are located directly below or within the VB of the material leading to some interaction with the VBM. The so called *p*–*d*-repulsion [84] leads to a slightly elevated energetic position of the VBM in the case of ZnO and CdO (Fig. 8). For the other compounds the *d*-states are located relatively deep below the VB leading to a decreased interaction strength with the VBM. In the case of In_2O_3 the *d*-states are energetically located already > 10 eV below the main VB and lead mainly to an interaction with the O-2*s* states. In the case of SnO_2 or GeO_2 the cation *d*-states are located even below the O-2*s* levels an are unlikely to have an influence on the VBM at all. For all compounds the top of the VB is mainly formed by non-bonding O-2*p*-orbitals, whereas the lower part of the VB is given by the bonding interaction between O-2*p* and M-2*s* (M = metal) orbitals. The CB is formed by their anti-bonding combination with a wide and dispersed free-electron-like lowest CB. The result of this particular electronic configuration are high ionization potentials (I_p) and electron affinities (χ_A) as well as small electron effective masses and high electron hole masses.

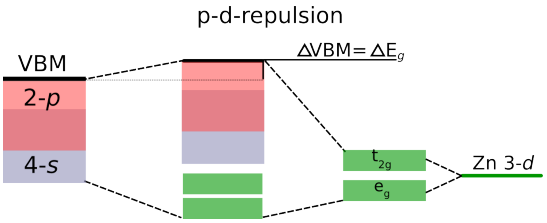


Figure 8: Interaction of shallow cation *d*-states with the VBM.

Ionization potential: The oxygen O-2*p* states have a low energetic position. As they form the VB and the cationic *d*-states are located well below the VB, this is not significantly modified. According to the common anion rule this leads to the high I_p (> 7.0 eV [85]) for all oxides with oxygen states at the VBM.

Note that in the case of copper based materials like Cu_2O (to the left of Zn) the occupied 3*d*-states already constitute the VBM instead of being located below the occupied oxygen states. This results in a position of the VBM at lower binding energies and smaller ionization potential [86]. The energetic position of *d*-states is decisive for the resulting materials properties. While *p*-type doping is difficult for the traditional TCOs, Cu_2O is known to be a *p*-type material [87].

Electron affinity: The reason for the energetically low CBM is the large cation *s*-orbital size. The resulting large overlap of those orbitals leads to a strong band dispersion. Therefore, the CB extends to relatively low energy values when compared with other covalent semiconductors or oxides with a CB formed by *d*-electrons. In fact, the electron affinity of many TCO materials reaches values as high as the ionization potential of silicon [88].

Effective masses: The large asymmetry in electron and electron hole effective masses can easily be understood in terms of the difference in spacial size of cation and anion orbitals, respectively. The band gap region of a typical TCO (In_2O_3) is shown in

Fig. 10 on the left. The highly dispersed CB is the fingerprint of all these materials in combination with a VB of higher effective mass.

Band gaps : The band gap has been debated for some of these oxides in the past. The most prominent examples are probably CdO and In_2O_3. The difficulties for CdO arised due the pronounced indirect nature of the band gap [89], which is unique among these oxides. In the case of In_2O_3, the VBM state at Γ is threefold degenerate and is derived from O-2p mixed with some In-4d (Γ_4, Tg symmetry), while the CBM state is a mixture of In-5s and O-2s orbitals (Γ_1, Ag symmetry). As the In_2O_3 crystal structure contains an inversion center and the electric-dipole operator is of odd parity, strong optical transitions are only permitted between two states of opposing parity. These symmetry requirements result in a vanishing optical transition matrix element

Table 2: Band gap trends for typical n-type TCO materials

	Group		
Period	X	XI	XII
3	3.3[a]	4.7[b]	5.3[c]
4	2.16[d]	2.6-3.1[e]	3.5[f]

a: Ref. [90] *b*: Ref. [91] *c*: Ref. [92]
d: Ref. [93] *e*: Ref. [94, 95, 96, 97] *f*: Ref. [98]

for direct VBM to CBM absorption. It is only from ~ 0.8 eV below the VBM that strong transitions are observed [99, 95].

Table 2 lists the band gaps of the common TCO materials. Generally, the energy gaps increase when going to the right in the periodic table and systematically decrease for the 4d row when compared with the 3d cations.

2.2.4 *Extrinsic substitutional doping*

The usual way to obtain conductivity in TCO materials is to use extrinsic doping. Table 3 lists the possible dopants for the most frequently used TCO host materials. The last column of the table contains the abbreviations which are in use for these materials. It is assumed that most of the listed dopants are substitutional impurities for the cation sites and selected from elements either one group to the right of the host cation or from the early transition metals with the appropriate valency.

The first option is the more traditional route and leads to the well known materials ITO, ATO and GZO. Doping with transition metals is becoming increasingly interesting as it can lead to improved electron mobility and optical properties [100]. The incorporated amount of dopants in the host materials is usually rather high ($> 10^{21}$ cm^{-3})

Table 3: List of *n*-type dopants for the three main TCO materials. The most frequently used combinations are underlined and their designations listed in the third column.

TCO	Dopant	Common designation
In_2O_3	Sn, Ti, Zr, Hf, F, Cl	ITO
SnO_2	Sb, F, Ta, Nb, W, Mo, Cl	ATO, FTO
ZnO	Al, Ga, In,B ,F, Cl	AZO, GZO, I ZO

when compared to typical intrinsic point defect concentrations. The actual content is, however, often not known. Fluorine, for example, can be used as anion replacing *n*-type dopant [101, 102] for all TCO hosts and is usually supplied in the form of HF during a CVD process. The actual amount of fluorine, which is captured by the sample, is then inferred from the electrical conductivity or free electron concentrations, which in many cases can substantially differ from the actual dopant's concentration. This is because the resulting free electron concentration depends on the concentration of incorporated *n*-dopant so that segregation effects and formation of competing phases (*e.g.* fluorites) diminish the active amount. Further, it is only the electronically active (ionized) impurities, which increase the number of free charge carriers. Finally, even in the case of fully ionized donors, the free electron density is only given by the uncompensated ionized impurities. For example in the case of ATO, a segregation has been conjectured to act as a limiting factor [103, 104], whereas in the case of ITO [105] but especially in AZO, a secondary phase formation may be relevant under certain conditions [106]. Interestingly, the ionization of these defects is usually not a problem at all. Fig. 9 shows the formation energy of the substitutional defects F_O and $Sb(V)_{Sn(IV)}$ calculated as a function of the Fermi energy using density functional theory [108].

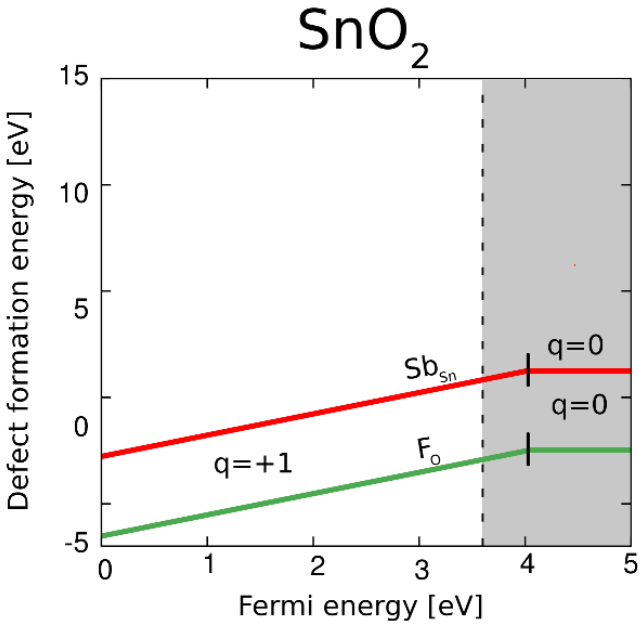


Figure 9: Defect formation energies of fluorine and antimony substitutional doping in SnO_2. The large range of existence of the charged defect state indicates its low or even negative ionization energy.

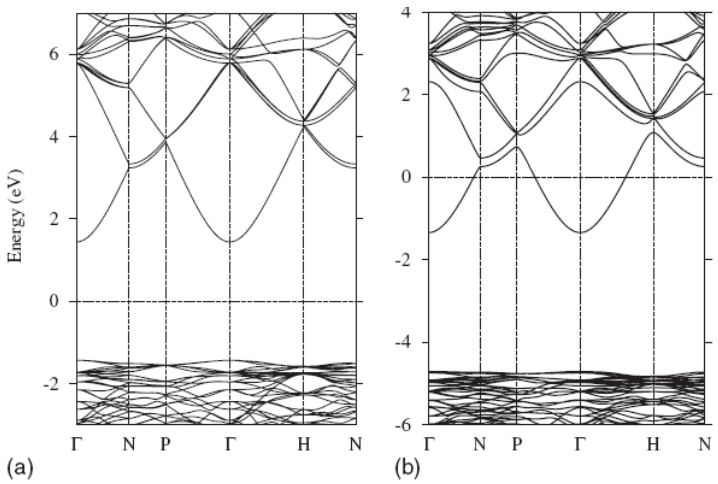


Figure 10: Band structure of pure In_2O_3 (a.) and In_2O_3 doped with Sn (b.) taken from Ref. [107]. The impact of the dopant is negligible and simply increases the Fermi energy into the band-gap.

Within the entire band gap both defects are ionized (q=+1) as indicated by the positive slope. A transition to a neutral charge state occurs only at a Fermi energy of $E_{CBM} + 0.5$ eV. This behavior indicates that the ionization energy with respect to the band edges is negative. For the frequently used host/dopant combinations such as In_2O_3:Sn, SnO_2:F, ZnO:Al, the defect related states of the dopant are located well above the CBM of the host or show a neat hybridization with the CB. In any case, no defect induced gap-states are produced by the dopants, but the defect states are resonances within the host states. The result is that the band-structures of the doped TCOs are almost unchanged even at relatively high dopant concentration on the order of $10^{20}cm^{-3}$. Fig. 10 compares the calculated band-structures of undoped (left) and 2.5% Sn-doped (right) In_2O_3 [107]. Apart from the location of the Fermi energy, the changes of the band-structure are marginal and consist of slight lifting in degeneracies at higher CB states. A different behavior is obtained when transition metal (TM) dopants such as Ti or Mo for In_2O_3 or Nb and Ta for SnO_2 are used. In those cases the TM can lead to occupied *d*-states within the band-gap in addition to resonances within the CB due to an exchange splitting of the *d*-states [100]. In essence, for usual doping procedures the ionization of the dopants within the TCO is not an issue, following the discussion above. Doping limits are nevertheless observed for all materials and caused by the thermodynamic stability of the materials as it will be discussed in the next section.

Taking In_2O_3 as an example, the effect of high doping concentrations has been studied in considerable detail. Frank and Köstlin have established a defect model for

Sn doped In_2O_3 [109]. At low concentrations Sn is incorporated in fully ionized form according to the reaction

$$Sn \leftrightharpoons Sn_{In}^{+} + e^{-} + In. \tag{2.5}$$

At higher doping concentrations additional oxygen is incorporated into the lattice on interstitial positions leading to neutral defect clusters of the kind

$$\frac{1}{2}O_2 + 2Sn_{In}^{+} + 2e^{-} \leftrightharpoons [2Sn_{In} - O_i]^0 \tag{2.6}$$

and

$$O_2 + 4Sn_{In}^{+} + 4e^{-} \leftrightharpoons [4Sn_{In} - 2O_i]^0, \tag{2.7}$$

which were termed reducible and non-reducible, respectively. The theory has further been elaborated thereafter using X-ray [110] and neutron diffraction [111] methods as well as atomistic [105] and electronic structure calculations [112]. The defect clusters as they were proposed by Frank and Köstlin could not be confirmed, but some association of acceptors with the dopants was found. The basic result nevertheless remained, which is that the doping limit in In_2O_3 is related to the incorporation of negatively charged oxygen interstitial defects.

2.2.5 *Intrinsic doping of TCOs*

While extrinsic doping is the primary means to achieve a high electrical conductivity in these materials, most TCOs already exhibit a remarkably high intrinsic conductivity under reducing conditions [20]. This conductivity is usually depending on temperature as well as oxygen partial pressure [113, 114, 115, 116, 117, 118, 119, 120, 121, 122]. As an example, Samson and Fonstad [114] measured the conductivity of SnO_2 samples for different temperatures and oxygen partial pressures (Fig. 11 (left)). Clearly, the conductivity increases with temperature and decreases with oxygen partial pressure. This effect strongly points to some intrinsic defect reaction such as oxygen vacancy or cation interstitial mechanism

$$0 \leftrightharpoons V_O^{2+} + \frac{1}{2}O_2 + 2e^{-} \tag{2.8}$$

$$Sn \leftrightharpoons Sn_i^{4+} + 4e^{-}. \tag{2.9}$$

Especially the electric conductivity is found to be proportional to the oxygen partial pressure as $\sigma_{el} \propto p^{-\frac{1}{6}}$, consistent with the oxygen vacancy model (V_O^{2+}) [114, 115, 116, 118]. In addition, the changes of electric conductivity are coupled to a non-stoichiometry, which was measured by gravimetric methods [123, 124]. In the case of SnO_2 two regimes of oxygen release from the sample were found. The low temperature signal was attributed to surface desorption of oxygen as well as surface reduction whereas the high temperature signal was assigned to the bulk reduction.

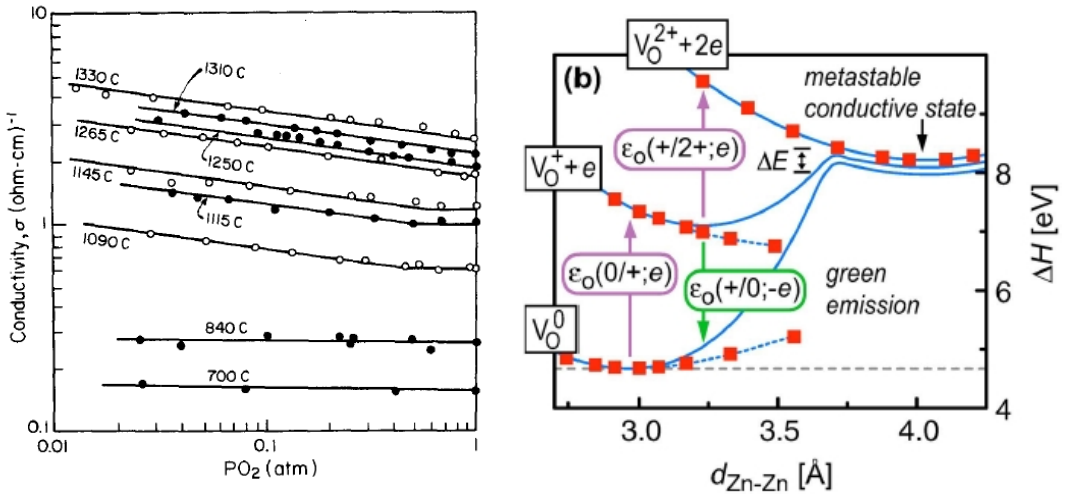


Figure 11: (left) The intrinsic conductivity as a function of oxygen partial pressure and at different temperatures (taken from Ref. [114])
(right) Energy as a function of the relaxation coordinate of the oxygen vacancy state in ZnO for different charge states. (taken from Ref. [125])

The nonstoichiometry of the bulk remained relatively low but measurable at higher temperatures ($c(V_O^{2+}) \sim 3.5 \times 10^{-5}$ at $p_{O_2} \sim 0.138$ atm and T= 1273 K).

Electronic structure calculations were conducted to reveal the nature of the intrinsic doping effects in SnO_2 [126, 127] and In_2O_3 [128, 129, 130, 131]. The results of these calculations are, however, not yet conclusive. For example, Kilic and Zunger proposed the existence of V_O^{2+} and Sn_i^{4+} as well as bound complexes of these two defects [126]. In contrast, Singh *et al.* reported that the vacancies in SnO_2 have high ionization energies and that generally all intrinsic donor defects have high formation energies (> 2.5 eV) [132]. As a result, no intrinsic conductivity or non stoichiometry should be possible. According to their calculations it is more likely that the conductivity is a result of hydrogen contamination. Early experiments already indicated that hydrogen and moist exposure can lead to the incorporation of hydrogen in the material [114, 133] with an additional donor level and increased conductivity. The hydrogen model, however, has difficulties in explaining the dependence on oxygen partial pressure and the background conductivity for cases where the presence of hydrogen can be excluded. A similar model was also proposed for In_2O_3 [134].

In contrast, Zunger and Lany [131] carried out first-principles defect calculations on In_2O_3, where they found deep oxygen vacancies but with low formation energies and with a peculiar meta-stable behavior. They argue that the intrinsic conductivity is caused by a meta-stable double photoexcitation of the oxygen vacancy as it is depicted in the energy configuration diagram in Fig. 11 on the right. In this model the deep vacancy state can be excited by light. The subsequent structural relaxation shifts the

vacancy state closer to the CB. Upon a second excitation the defect state shifts even further and forms a resonant state within the CB. Because the vacancy state is not deep (an electron trap) in this ionized state, electrons remain within the CB. This model of persistent photoconductivity was originally proposed for ZnO [125], where Lany and Zunger reported a vacancy state even deeper within the band gap compared to In_2O_3.

2.2.6 *TCO performance*

For many applications it is desirable to optimize both, the conductivity and the transparency of the TCO material. This is difficult, because the two properties compete with each other in several ways. On the one hand transparency requires the existence of a large band gap (> 3.0 eV). On the other hand the carrier mobility is a function of the band gap according to the $\mathbf{k} \cdot \mathbf{p}$ theory [135]. Only small band gap materials have also small effective masses and consequently a large electron / electron hole mobility. More importantly, the doping ability, *i.e.*, the maximum free charge carrier density of a semiconductor or insulator scales with the fundamental band gap. The larger the band gap, the higher the likelihood for a so-called Fermi-pinning level. For all materials the range of Fermi levels is confined within an energy range, largely independent of the size of the band gap. In certain cases this restriction of Fermi level is universal and not even very much material dependent. This situation, for example, is often encountered under extensive hydrogen contamination. The Fermi energy range is determined by the properties of hydrogen independent of the material [136].

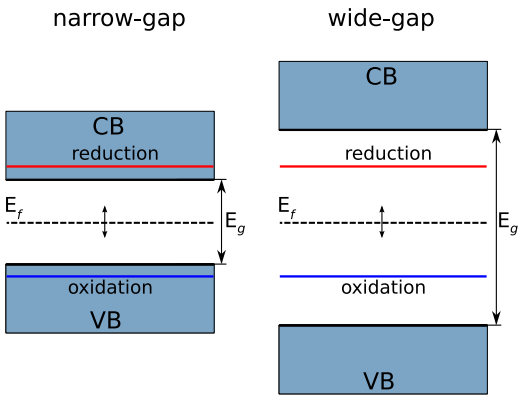


Figure 12: Illustration of Fermi pinning levels.

In order to obtain free charge carriers which contribute to electronic transport the Fermi level has to be close or within either the VB or CB. Three different classes of semiconductors can be distinguished:

- The Fermi level can enter both VB and CB
- The Fermi level can enter either VB or CB
- The Fermi level cannot enter either of the bands

The width of the band gap as well as the energetic position of the band edges with respect to the vacuum level determine to which class a semiconductor belongs. Fig. 12 illustrates schematically this so-called Fermi pinning behavior. The microscopic

origin of the pinning is usually associated with redox processes within the material. Frequently, these redox processes are manifested in the formation of point defects, which effectively compensate the free holes (oxidation) or electrons (reduction). In this light, establishing a TC based on an insulating material makes the use of insulators with only a moderately large band gap necessary for achieving a dopability with at least one carrier type. In addition to a low band gap, it is required that either χ_A is high or I_p is low in order to facilitate the generation of free electrons or holes, respectively.

Several additional effects exist, which can lead to the deterioration of the TCO performance. Doping (*e.g.* n-type) necessarily increases the free electron concentration within the CB. An increased free electron density in turn lifts the plasma frequency,

$$\omega_p = \sqrt{\frac{n_e e^2}{m_e^* \epsilon_0}}, \tag{2.10}$$

where ϵ_0 is the vacuum permittivity and e the elementary charge. Fig. 13 shows the transmission spectrum of pure and antimony doped SnO_2 for two different doping concentrations. The transparency window is narrowed by increased doping, which can be beneficial for some applications (low-e glass), but undesirable for applications, where a good transparency is needed in the low frequency regime, like solar cell transparent contacts. A second effect of the heavy doping influences the carrier mobility. An increased doping concentration leads to an increased number of ionized impurities, which in turn cause a decay of the carrier mobility [137]. Finally, when the CB is heavily occupied by carriers, excitations into higher CB states become possible (Fig. 14). This effect may lead to a decay of the transparency in the high frequency regime. Only materials with a sufficiently wide (dispersed) CB can maintain high frequency transparency coexisting with a high free electron concentration.

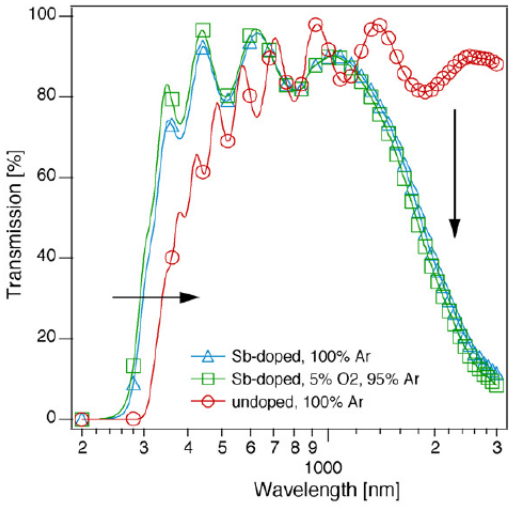


Figure 13: Transmission spectrum of doped and undoped SnO_2 taken from Ref. [138]

In terms of transparency, occupation of higher states within the CB can also have a beneficial effect as it is shown in Fig. 14. Because the bottom part of the CB is heavily occupied, excitations between the VBM and CBM are no longer allowed. Initially, this effect leads to an increase in optical band gap as a function of carrier concentration and referred to as the Burstein-Moss (BM) shift [139]. Thus, free carrier concentration, carrier

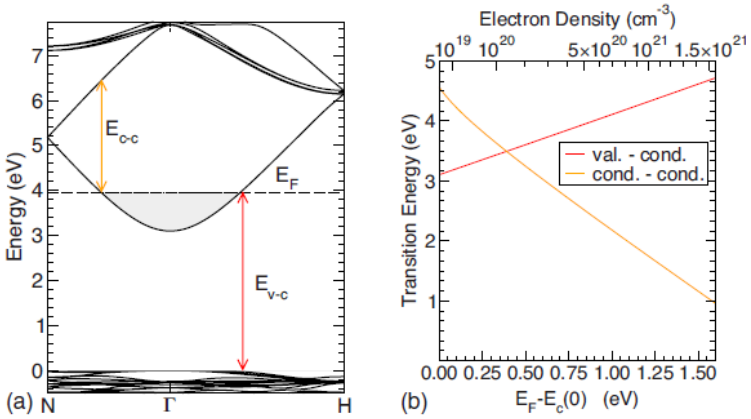


Figure 14: (left) Schematic illustration of VB to CB and CB to CB excitations. (right) Absorption onsets of the two mechanisms as a function of carrier concentration (from Ref. [97]).

mobility and optical transmission of these wide-gap semiconductors are strongly interrelated and have to be optimized with respect to each other. For a good electric conductivity, relatively large carrier densities are necessary. As a general rule it is, however, advisable to increase the carrier mobility instead of the free carrier concentration [140], because this does not increase the plasma frequency. On the other side, the carrier mobilities are bound to the intrinsic limitations imposed by the effective mass of the host bands, but especially to ionized impurity and other scattering mechanisms. Consequently, it is desirable to maximize the number of free charge carriers per ionized impurity for a good mobility.

Besides these major issues, several other factors are relevant when selecting a TCO for a specific application. Most important criteria together with the best presently available TCO materials are presented in Table 4. In addition to the afore mentioned properties, the durability, etchability, and resistance to hydrogen plasmas as well as the hardness are listed. Currently, issues like the price or abundance of the base materials are considered in addition, because some technologies are expected to rapidly grow (*e.g.* solar cell technologies). The listing in Table 4 is adopted from Gordon [6] and should be understood as a rough guide within the large variety of TCs.

2.2.7 *Availability of materials*

In many research articles as well as conference contributions a low abundance of especially indium is put forward as a motivation for TCO research. Because ITO is in many ways an ideal material [20, 21, 22] especially for the display industry, shortages and the volatility of the indium metal price is a major concern and leads to the search

Table 4: List of *best* TCOs for different applications. (adopted from Ref. [6])

Property	TCO
Highest transparency	ZnO:F, Cd_2SnO_4
Highest conductivity	In_2O_3
Lowest plasma frequency	SnO_2,ZnO:F
Highest plasma frequency	Ag, TiN, In_2O_3:Sn
Highest work function, best contact to p-Si	SnO_2:F, $ZnSnO_3$
Lowest work function, best contact to n-Si	ZnO:F
Best thermal stability	SnO_2:F, TiN, Cd_2SnO_4
Best mechanical durability	TiN, SnO_2:F
Best chemical durability	SnO_2
Easiest to etch	ZnO:F, TiN
Best resistance to H plasmas	ZnO:F
Lowest deposition temperature	In_2O_3, ZnO:B, Ag
Least toxic	ZnO:F, SnO_2:F
Lowest cost	SnO_2:F
Most abundant	ZnO:Al

for ITO replacements. Fig. 15 shows the relative abundance of the elements in the earth crust on a logarithmic scale (Ref. [141]) and the traditional cation species used for TCOs are highlighted in red. Clearly, indium metal is the most scarce of these elements along with cadmium. It is, however, interesting to compare the abundance of In with that of silver which is about equal. In the case of indium, the world production amounts to approximately 586 tons per year (2008) [142], whereas it is on the order of 20,000 tons per year (2008) for silver [143]. In the latter case no sign of shortages are perceivable. Nevertheless, China, the world largest indium producing country runs indium refineries with 75% of their capacity [144] leading in fact to shortages in primary indium supply. The reason is the dependence of indium refinement on so called indium bearing concentrates which are obtained as a byproduct of mainly zinc production [145]. Due to the relatively low worldwide indium consumption and the high price volatility, only very few zinc producers bother to extract indium at all. In the meantime large FPD producing companies in Asia have launched efforts to recycle indium. Especially, the sputter deposition methods which are mainly employed make only use of about one third of the ITO targets. In the past, shortages of indium could easily be compensated by recycling.

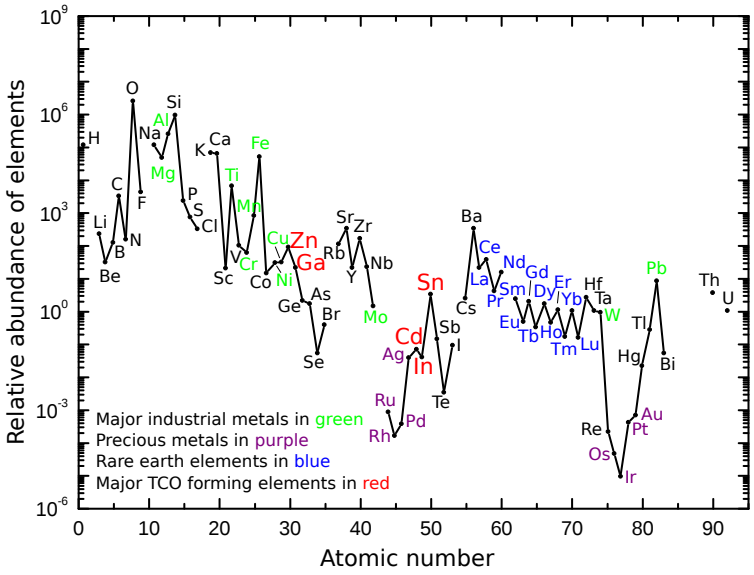


Figure 15: Relative abundance of elements in the earth crust (Ref. [141]) with Si being normalized to a 10^6. Different groups of materials are highlighted.

2.3 OPEN QUESTIONS

Following the discussion of the previous chapter, several important questions can be formulated. All topics listed below are addressed in subsequent chapters of this thesis.

2.3.1 *Intrinsic defects*

The mechanisms of extrinsic doping appear to be well understood, while the intrinsic properties are controversial. The defect chemistry of all oxides has been studied for a considerable time. Nevertheless, the relevant defect equilibria in In_2O_3, SnO_2 and ZnO are not satisfactorily known yet. From a modeling point of view this is due to the limitation of the standard density functional theory methodology to access defect properties. A recalculation of the defect energetics using an improved methodology is necessary at this stage.

2.3.2 *Doping / doping limits*

The doping ability of a material is restricted by Fermi pinning processes, which in the case of extrinsic *n*-doping involve the occurrence of intrinsic acceptor defects. According to presently published calculations [131, 126], the doping limits of TCOs like SnO_2 and In_2O_3 seem severely underestimated. Again, this is due to limitations

of standard DFT methodology. For In_2O_3 the doping limit is fairly well known from experimental studies [109]. In SnO_2 this is not the case and the identification of the primary acceptor as well as a realistic estimation of a doping limit is necessary to map out the potential of doped SnO_2 based materials as a competitor with ITO.

2.3.3 *p-type doping in conventional TCOs*

By the same token the limits toward a low Fermi level are presently of interest. Previous DFT calculations predict that *p*-type doping is possible in some TCO materials [132, 146]. In contrast, experiments are controversial. Using an improved DFT methodology a realistic estimation of *p*-type doping ability is needed.

2.3.4 *Kinetics*

The kinetic properties of point-defects are a largely unknown area for almost all TCOs . While diffusion experiments have been undertaken in the case of In_2O_3 and SnO_2, the range of environmental conditions is usually limited in experiments and the doping conditions of the material is often not well specified. The prediction of the defect mobilities using first-principles method is a reliable alternative and can contribute to the interpretation of experiments.

2.3.5 *Entropy*

The point defects in TCOs have only been studied in the zero temperature approximation, *i.e.*, by neglecting entropy contributions of defect formation. In principle, both the formation and the migration parameters of point-defects can be significantly temperature dependent. Particularly, in the light of the controversial issue of intrinsic defect equilibria the magnitude of such contributions needs to be estimated.

2.3.6 *Surfaces*

A large body of data on TCOs electronic structure was obtained via photoelectron techniques (XPS/UPS) [51, 85] while the atomic structure of the surfaces is not well understood for many of the TCOs. In contrast to In_2O_3, SnO_2 and ZnO have been studied in more detail. Still, in the case of SnO_2, the (110) surfaces remained an unresolved problem despite numerous surface studies. In the case of In_2O_3, almost no atomistic data is available, despite its importance for electric contacts and gas sensing.

Part II

METHODS AND FUNDAMENTALS

In the following, fundamental equations of density functional theory are reviewed. The most important concepts such as exchange-correlation functionals, basis-sets and pseudo-potentials are briefly introduced. In the second part the formalism and important aspects of first-principles thermodynamic calculations on point defects and surfaces are described.

3

DENSITY-FUNCTIONAL THEORY

3.1 HOHENBERG KOHN THEOREM

In the following a brief overview on central topics of density functional theory (DFT) calculations is given. Comprehensive review articles and books are available describing different aspects of from fundamentals to applications [147, 148, 149].

Starting point for any quantum-mechanical treatment of a system is the well established Schrödinger equation $\hat{H}\Psi = E\Psi$ with

$$\begin{aligned} \hat{H} &= -\frac{\hbar^2}{2m_e}\sum_i \boldsymbol{\nabla_i}^2 - \sum_{i,I}\frac{Z_I e^2}{\mid \mathbf{r}_i - \mathbf{R}_I \mid} + \frac{1}{2}\sum_{i\neq j}\frac{e^2}{\mid \mathbf{r}_i - \mathbf{r}_j \mid} \\ &- \sum_I \frac{\hbar^2}{2M_I}\boldsymbol{\nabla_i}^2 + \frac{1}{2}\sum_{I\neq J}\frac{Z_I Z_J e^2}{\mid \mathbf{R}_I - \mathbf{R}_J \mid} \end{aligned} \tag{3.1}$$

and the many body wave function Ψ. The Hamiltonian ($\hat{H}$) includes the kinetic energies of all electrons and nuclei (terms 1 and 4 of Eq. 3.1) and the Coulomb electron-electron, electron-nucleus and nucleus-nucleus interactions (terms 2, 3 and 5 of Eq. 3.1)*. Upper case letters R_I and M_I refer to nuclear coordinates and mass, whereas electron coordinates and mass are given by r_i, m_e and electron and nuclear charge by e and Ze, respectively. Neglecting the fourth term is the Born-Oppenheimer approximation and valid for most solids due to the higher mass of nuclei compared to electrons. The most difficult terms to treat within this differential equation are the electron-electron interactions, because electrons contribute a large portion to the total energy. In addition, the accuracy which is necessary in order to be useful in practical calculations (*e.g.* a fraction of a bond energy) is high when compared with the typical magnitude of these electron electron interactions and thus good approximations are needed for this so called many-body problem.

Conceptually, two methods have emerged which deal with this problem in an efficient and reasonably accurate way: Hartree-Fock theory and density-functional theory. In both cases the paradigm is to abandon the many electron wave function ($\Psi(\mathbf{r_1}, \mathbf{r_2}, ..., \mathbf{r_n})$) and introduce simpler objects to deal with instead. By solving the

*Within this chapter Hartree atomic units are used: $\hbar = m_e = 4m/\epsilon_0 = 1$

respective Schrödinger like equations, where $\hat{H}$ is replaced by a more simple mean field Hamiltonian ($\hat{H}^{DFT}$ or $\hat{H}^{HF}$), one obtains a series of single particle eigenstates (ϕ_i) and their corresponding eigenenergies (ϵ_i). In the case of DFT, this simpler object is the electron density $n(\mathbf{r})$, which contains only one set of coordinates. The foundation for the theory is contained in the original papers of Hohenberg, Kohn and Sham [150, 151]. The two Hohenberg-Kohn theorems state that

- There is one and only one external potential V_{ext} (*e.g.* ionic cores or electric fields) corresponding to a given ground state electron density,
- It is only the ground state electron density which minimizes the total energy of the N electron system.

Since the external potential $V_{ext}(\mathbf{r})$ also determines the wave function, the wave function itself is implicitly but uniquely defined by the ground state electron density. Therefore, a universal energy functional ($F[n(\mathbf{r})]$) exists which depends on the electron density and which, together with the system dependent energy contribution due to the external potential $\int V_{ext}(\mathbf{r})n(\mathbf{r})d\mathbf{r}$, minimizes the total energy $E[n(\mathbf{r})]$ when the ground state electron density $n_0(\mathbf{r})$ is given as the argument,

$$E_0 = \langle \Psi_0 \mid \hat{H} \mid \Psi_0 \rangle = F[n_0(\mathbf{r})] + \int V_{ext}(\mathbf{r})n_0(\mathbf{r})d\mathbf{r}. \tag{3.2}$$

This is analogous to the variational principle of quantum mechanics which carries over to the electron density in the case of DFT. There is no unique prescription how the universal functional $F[n(\mathbf{r})]$ is to be constructed. The almost exclusively employed Kohn-Sham (KS) approach goes as follows. First, the universal functional is divided into parts which from a physical viewpoint are contributions to the total energy

$$F[n(\mathbf{r})] = T^S[n(\mathbf{n})] + \frac{e^2}{2}\int\int \frac{n(\mathbf{r})n(\mathbf{n}')}{\mid \mathbf{r}-\mathbf{r}' \mid} d\mathbf{r}d\mathbf{r}' + E^{XC}[n(\mathbf{r})]. \tag{3.3}$$

The first term represents the kinetic energy, whereas the second term called the Hartree term describes the Coulomb potential energy of a classical charge distribution $n(\mathbf{r})$. All remaining contributions to the energy are grouped together in the last term, which is called the exchange correlation (XC) energy functional. Only for the Hartree energy an explicit expression can be given. However, from quantum mechanics it is known that the kinetic energy is a very large portion of the total energy and should be approximated fairly well. The beauty of the KS scheme is that in order to evaluate the kinetic energy functional in Eq.3.3 the electron density is expanded in some fictional single particle orbitals ϕ_i

$$n(\mathbf{r}) = \sum_i f_i \mid \phi_i(\mathbf{r}) \mid^2 \tag{3.4}$$

with occupation numbers f_i. This way the kinetic energy can elegantly be expressed in the usual way

$$T^S[n(\mathbf{r})] = \sum_i f_i \langle \phi_i \mid -\frac{\hbar^2}{2m_e}\boldsymbol{\nabla}_{\mathbf{r}}{}^2 \mid \phi_i \rangle. \tag{3.5}$$

Finally, the XC term has to be approximated in some way. This last term is a relatively small contribution to the total energy, which even makes crude approximations like the local density approximation (LDA) very successful [152]. The ground state energy and ground state electron density are obtained by applying variations to the density

$$\delta n(\mathbf{r}) = \delta\phi_i^*(\mathbf{r})\phi_i(\mathbf{r}) \tag{3.6}$$

with the constraints of conservation of particle number and orthonormality of the KS orbitals

$$\int n(\mathbf{r})d\mathbf{r} = N \quad \text{and} \quad \int \phi_i^*(\mathbf{r})\phi_j(\mathbf{r})d\mathbf{r} = \delta_{ij}, \tag{3.7}$$

expressed by Lagrange multipliers μ and ϵ_i. The minimization of the total energy with respect to variations of the electron density results in the Schrödinger like KS equations

$$\left[-\frac{\hbar^2}{2m_e}\nabla_{\mathbf{r}}^2 + V_{eff}(n(\mathbf{r}))\right]\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}) \tag{3.8}$$

with

$$V_{eff}(\mathbf{r}, n(\mathbf{r})) = V_{ext}(\mathbf{r}) + e^2\int \frac{n(\mathbf{r}')}{\mid \mathbf{r}-\mathbf{r}' \mid}d\mathbf{r}' + \frac{\delta E^{XC}[n(\mathbf{r})]}{\delta n(\mathbf{r})}. \tag{3.9}$$

Equations 3.8 and 3.9 are formally very similar to Schrödinger's equation and to the Hartree-Fock equation. In fact, the important difference to Hartree-Fock theory is the replacement of the exchange energy by the unknown XC-potential $\delta E^{XC}[n(\mathbf{r})]/\delta n(\mathbf{r})$.

The KS Hamiltonian depends on the KS eigensolutions or similarly on the calculated electron density. Therefore, like in the Hartree-Fock theory Eq. 3.8 can be solved alternative to a direct minimization of the energy functional by the self consistency criterion. Most advanced codes like *e.g.* **vasp** [153, 154] or **ABINIT** [155] solve the KS equations in the following way.

The external potential of the ion cores is constructed by the superposition of either the Coulomb potentials or so called pseudo-potentials (PP). An initial charge density or set of KS orbitals is selected. This can either be a completely homogeneous charge density or a "better" guess based on earlier calculations. For the representation of the orbitals and density, a suitable basis set is selected (see Sect. 3.3). The Hartree and XC-potentials are calculated in order to set up the initial Hamiltonian of the system. The Hamiltonian is applied to the trial KS orbitals. Using an iterative matrix diagonalization procedure, the (at least) N lowest eigenvalues and eigenvectors of the system are determined. The new eigenvectors can be used to calculate a new electron density and construct a new Hamiltonian. This self-consistent cycle is usually terminated when the total energy difference compared to the previous cycle falls below a predefined limit. In any new cycle, the new charge density can be used as input, but is usually mixed with the old one to improve the convergence.

3.2 EXCHANGE-CORRELATION FUNCTIONALS

3.2.1 *(Semi-)local functionals*

Kohn and Sham already pointed out that solids can often be considered in the limit of the homogeneous electron gas. Although exchange and correlation effects are inherently of non-local character, they are local in the limit of a homogeneous electron gas. In that case, the exchange correlation energy density at each point can taken to be the same as in a homogeneous electron gas with that density:

$$E_{XC}^{LDA}[n] = \int n(\mathbf{r})\epsilon_{hom}^{xc}[n(\mathbf{r})]d\mathbf{r} = \int n(\mathbf{r})\epsilon_{hom}^{x}[n(\mathbf{r})] + n(\mathbf{r})\epsilon_{hom}^{c}[n(\mathbf{r})]d\mathbf{r}. \tag{3.10}$$

The information needed is the exchange and correlation energy of the homogeneous electron gas as a function of the electron density. The exchange part E^X is given by the simple analytic form [156]:

$$\epsilon_{hom}^{X}(n) = -\frac{3}{4}\left(\frac{3}{\pi}\right)^{\frac{1}{3}} n(\mathbf{r})^{\frac{4}{3}}. \tag{3.11}$$

The exchange energy describes the energy gain due to the requirement of anti-symmetric wave functions for fermions. This interaction is purely of quantum mechanical origin and results in a larger average distance between electrons, which on the other hand reduces the mutual Coulomb repulsion between electrons. This exclusion of electrons of the same spin type is often referred to as exchange hole. The shape of the exchange hole is modified to some extent by correlation effects (correlation hole). The contribution of correlation (E^C) is usually small, and was computed with high numerical accuracy for the case of the ideal homogeneous electron gas using Quantum Monte-Carlo methods [152]. Within the LDA and its generalization LSDA (local spin density approximation), it is assumed that effects of correlation and exchange are short ranged. This approximation works best for highly delocalized systems such as metals.

Beyond the LDA, the gradient of the exchange-correlation energy density at each point could be considered within the calculation. Such a straightforward approach did not lead to improvements in the past, because sum rules and various other conditions were violated [157]. In GGA (generalized gradient approximation) the XC-gradient enters the energy functional in a more complex way:

$$E_{XC}^{GGA}[n] = \int n(\mathbf{r})\epsilon_{xc}(n, \nabla n)d\mathbf{r} = \int n(\mathbf{r})\epsilon_{x}^{hom}(n)F_{xc}(n, \nabla n)d\mathbf{r}. \tag{3.12}$$

The most widely applied functionals were proposed by Perdew, Wang and Ernzerhof (PBE) [158] Becke (B88) [159] Perdew Wang (PW91) [160] which altogether show a very similar performance. The LDA usually favors "overbinding" that results in too high cohesive energies and too small lattice constants and bond lengths [161]. The GGA usually improves on this by softening bonds and decreasing cohesive energies, but there is no general tendency with respect to experimental data.

3.2.2 *Orbital-dependent functionals*

One class of exchange correlation functionals which deserve a special attention are the so called hybrid functionals. The idea behind these functionals is to incorporate some non-local Hartree-Fock exchange into either an LDA or GGA functional.

Becke's B3LYP functional [162] is popular in quantum chemistry, where it allows to obtain energies with almost chemical accuracy. In the case of solid systems, and especially metals, this functional has, however, failed because [163] it is adjusted to describe mainly molecular properties. In addition, the calculation of the Hartree-Fock exchange energy

$$E_X^{HF} = -\frac{e}{2}\sum_{\mathbf{k}n,\mathbf{q}m} f_{\mathbf{k}n} f_{\mathbf{q}m} \times \int\int d^3\mathbf{r} d^3\mathbf{r}' \frac{\phi_{\mathbf{k}n}(\mathbf{r})\phi_{\mathbf{q}m}(\mathbf{r}')\phi_{\mathbf{k}n}(\mathbf{r}')\phi_{\mathbf{q}m}(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} \tag{3.13}$$

is computationally expensive with a plane wave basis set commonly used in the solid state community. Only recently a new type of screened hybrid functional was proposed by Heyd and coworkers. In this HSE [164, 165, 166] scheme the GGA-PBE [158] exchange energy is combined with Hartree-Fock exchange in the following way

$$E_{XC}^{HSE} = \frac{1}{4}E_x^{HF,SR}(\mu) + \frac{3}{4}E_x^{PBE,SR}(\mu) + \frac{3}{4} + E_x^{PBE,LR}(\mu) + E_c^{PBE}, \tag{3.14}$$

where SR and LR denote the short range and long range part and μ is a screening length which defines the range separation between SR and LR. The screening is typically modelled by an error function and the Coulomb kernel is decomposed into the form

$$\frac{1}{r} = SR_\mu(r) + LR_\mu(r) = \frac{erfc(\mu r)}{r} + \frac{erf(\mu r)}{r}. \tag{3.15}$$

Depending on the the range-separation parameter, the following functionals can be distinguished:

$$\mu = 0 \rightarrow \text{PBE0}, \quad \mu = 0.2 \rightarrow \text{HSE06}, \quad \mu = \infty \rightarrow \text{GGA-PBE}.$$

In practice, the range separation and the mixing parameters (0.25 in the previous equations) are often adjusted, because there exists no unique combination for all materials. An alternative approach is to remedy the self-interaction problem which persists in LDA and GGA functionals by adding an orbital dependent and atom specific repulsive term for certain localized f- or d-states. This approach is termed the LDA/GGA+U [167, 168] where the U is analogous to the electron-electron repulsion in the Hubbard model [169, 169]. Also this parameter is adjusted in order to fix certain LDA/GGA deficiencies.

3.3 BASIS SETS AND PSEUDO-POTENTIALS

In *ab initio* calculations the basis set is a choice of convenience, system type, accuracy and speed. A very intuitive expansion for the wave function is given by localized

atomic orbitals. The wave functions can be represented by a linear combination of Gaussians (GTO), Slater type (STO) or numerical radial atomic(-like) orbitals [161]. Although the calculations are often very efficient and the primary choice in "order-N" methods [170], there are, however, some major drawbacks regarding their transferability. The orbitals must be chosen specifically for a given system to be accurate and efficient [171]. A more systematic basis set is given either by the real-space basis or the plane wave basis, because they are orthonormal and complete so that any arbitrary function is representable without prior knowledge. Especially for spatially extended wave functions, like in solid crystalline systems, plane waves are a natural choice.

Bloch's theorem states that the wave function in a crystal is the product of a periodic function u_n and a phase factor [61],

$$\psi_{n,k} = u_n \exp(i\mathbf{k} \cdot \mathbf{r}), \tag{3.16}$$

where $\mathbf{k}$ is a wave vector within the first Brillouin-zone. By introducing a plane wave expansion for u_n, an infinite sum over lattice vectors $\mathbf{G}$ is obtained for each single particle wave function

$$\psi_{n,\mathbf{k}} = \sum_{\mathbf{G}} c_{n,\mathbf{k}+\mathbf{G}} \exp[i(\mathbf{k}+\mathbf{G}) \cdot \mathbf{r}]. \tag{3.17}$$

The sum is truncated at some highest $\mathbf{G}$ and usually given in terms of a kinetic energy cutoff

$$E_{cut} = \frac{(\mathbf{k}+\mathbf{G})^2 \hbar^2}{2m_e}, \tag{3.18}$$

so that the space resolution of the basis-set is independent of system volume for a given E_{cut}. The major strength of the plane wave basis is the efficient evaluation of the kinetic energy (Eq. 3.18) and the electrostatic energies in Fourier space. The drawback is that the size of the basis increases rapidly with cell volume (even for vacuum). Secondly, the rapidly oscillating wave functions in the core region make a large number of plane waves necessary. In practice this problem is circumvented by using effective core potentials, the so called pseudo-potentials (PP).

PPs combine the potential due to the core electrons and the nucleus in an effective potential. This approach is justified, because the localized core electrons cannot overlap with nearby atoms and have, therefore, no influence on bonding. The norm-conserving ab initio PP, for example, is constructed in such a way that a pseudo wave function has the same value as the true wave function outside a cut-off radius r_c around the core. A list of additional requirements for a "good" ab initio PP was given by Hamann, Schlüter and Chiang [172]. The all-electron and pseudo valence eigenvalues should agree for the chosen atomic reference configuration and the logarithmic derivatives of the all-electron and pseudo wave functions should match at r_c. In addition the integrated charge inside r_c for all-electron and pseudo wave functions should be equal (norm-conservation) [172]. The goal is to obtain smooth PPs in order to minimize the range of Fourier components needed to describe the valence properties to a given

accuracy. By increasing the cut-off radius, for example, a greater part of the strongly oscillating wave function within the core region is avoided, and replaced by a smooth function. If the above mentioned criteria are fulfilled one has some freedom to obtain the smoothest function possible to represent the core region [173].

Most recent approaches drop the requirement for norm conservation in order to retain more smoothness. These so called "ultra-soft" PPs proposed by Vanderbilt [174] keep the accuracy by transformation of the problem in terms of a smooth function and an auxiliary function around each ion core that represents the rapidly varying part of the density.

A related, but more general method is the projector augmented wave (PAW) method by Blöchl [175] which also employs localized auxiliary functions. The difference is that all information on the core wave function is retained similar to linear augmented plane wave methods [176]. The PAW method strictly is an all electron approach, where the core part of the wave function is frozen.

4

THERMODYNAMICS AND KINETICS

4.1 THE SUPERCELL APPROACH

A supercell describes a multiplication of the unit cell along its crystallographic axes, with periodic boundaries applied in all directions. This enlarged unit cell serves as the system in a thermodynamic sense and describes a perfect uniform solid. Within this system, defects can be introduced and the change in total energy and other quantities can be calculated in principle. In practice, the system size is limited to a few hundred atoms by computational resources so that this single phase region without any exterior can be simulated only. With this restriction, two main routes can be followed. Calculations of the defect energetics can be conducted within a canonical ensemble. The second method is to drop the constraints of constant particle number and charge and instead, introduce the grand-canonical ensemble to describe the defect energetics. This second route is used mainly in this study, but the first method is also employed in some cases.

4.2 POINT DEFECT THERMODYNAMICS

For a material in equilibrium with reservoirs of entropy and particles, the grand-canonical ensemble is most appropriate for the calculation of the equilibrium concentration of defects. The isobaric grand-potential of a multicomponent system is defined as

$$0 \equiv \Phi_G = E - TS + PV - \sum_i \mu_i n_i - \mu_e n_e \tag{4.1}$$

where the pressure (P), the temperature (T) and the chemical potentials (μ_i) as well as the electrochemical potential (μ_e) (Fermi energy) are the control variables and E, S, V and n_i the total energy, entropy, volume, and particle numbers, respectively.

The process of point defect creation is pictured in Fig. 16, including exchange of atoms and electrons with their reservoirs. The figure indicates that the material in question can be in equilibrium with other phases such as oxygen gas or metal in the case of a metal oxide. By separating the configurational entropy of point defects, the

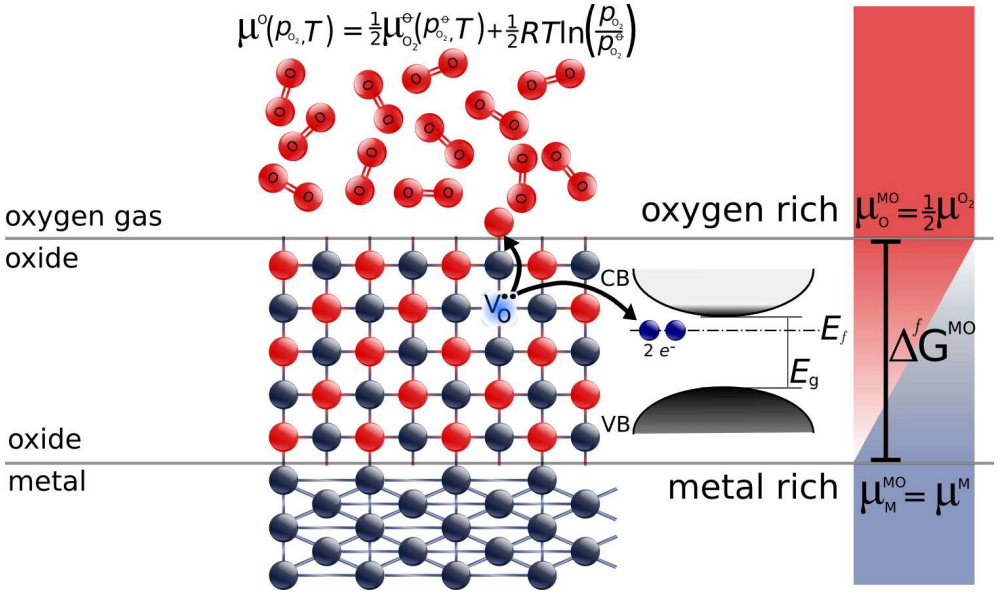


Figure 16: Defective metal oxide and the potential reservoirs for anions, cations and electrons.

equilibrium state of the system is achieved for fixed $\mu_1, \mu_2, ..., E_f, T, P$

$$T(S^{def}_{\text{conf}} - S^{id}_{\text{conf}}) = \underbrace{E^{def} - E^{id}}_{\Delta E_f} - T\underbrace{(S^{def}_{\text{other}} - S^{id})}_{\Delta S_f} + P\underbrace{(V^{def} - V^{id})}_{\Delta V_f} \tag{4.2}$$

$$- \sum_i \mu_i \underbrace{(n^{def}_i - n^{id}_i)}_{\Delta n_i} - E_f \underbrace{(n^{def}_e - n^{id}_e)}_{q} \tag{4.3}$$

where ΔE_f, ΔS_f, ΔV_f are the defect excess energy, entropy and volume and Δn_i, q describe the number of exchanged particles upon defect creation. Assuming that only few defects ($n < 1\%$) populate a very large lattice (N sites), the configurational entropy can be explicitly formulated as a function of the defect concentration c_D

$$\Delta\Omega = k_B T \ln \binom{n}{N-n} \approx k_B T \ln c_D = \Delta E_f - T\Delta S_f + P\Delta V_f - \sum_i \mu_i \Delta n_i + E_f q \tag{4.4}$$

where $\Delta\Omega_f$ is the grand canonical defect formation energy. For simplicity, it will be referred to as the defect formation energy in the following. In the case of charge neutral and stoichiometric defect pairs such as Frenkel and Schottky defects, the last two terms vanish and the quantity $\Delta\Omega_f$ is equivalent to the more familiar Gibbs energy of defect formation ΔG_f.

4.2.1 *Chemical reservoirs*

The final two quantities in Eq. 4.4 describe the reservoirs for particle and electron exchange. The chemical potentials of the atomic species in a non-elemental crystal are inherently troublesome quantities to measure, since it is not possible without further information to assign specific fractions of the total free energy of the crystal to the various species present. In reality, the chemical potentials are set by the experimental conditions, which define the properties of the reservoirs of atoms in equilibrium with the sample. The same is true for the electrochemical potential (final term in Eq. 4.4), which is determined by the *e.g.* doping condition of the material.

If, for example, an oxide phase is annealed in an oxygen atmosphere, the chemical potential of the oxygen atoms is the same as in gaseous molecular oxygen at the experimental partial pressure and temperature. For ideal gases, the temperature and pressure dependence can be expressed via the ideal gas law [177] as

$$\mu_{\mathrm{O}}(p_{\mathrm{O}_2}, T) = \frac{1}{2}\left\{\mu_{\mathrm{O}_2}(p^{\circ}_{\mathrm{O}_2}, T) + k_B T \ln\left(\frac{p_{\mathrm{O}_2}}{p^{\circ}}\right)\right\}. \tag{4.5}$$

Considering the specific case of a metal oxide ($\mathrm{M}_x\mathrm{O}_y$) the question arises of how the chemical potential of the metal species (μ_{M}) can be fixed. Such a control could in principle be established in experiment using an alloy as a substrate which contains the metal within the oxide as alloying element. In such metallurgical systems the chemical potential of the metal is determined by the activity of the metal a_{M}, which is connected to the solute concentration via the activity coefficient $a_i = \gamma_i x_i$,

$$\mu_{\mathrm{M}}(T) = \mu^{\circ}_{\mathrm{M}} + k_B T \ln\left(a_{\mathrm{M}}\right) \approx \mu^{\circ}_{\mathrm{M}} + k_B T \ln\left(x_{\mathrm{M}}\right). \tag{4.6}$$

For an ideal solid solution, the activity coefficient is unity. There are also cases (*e.g.* Zn) where control can be established via the gas pressure of the metal vapor, so that Eq. 4.5 could be applied, again. If a full equilibrium is justified and the considered compound system is sufficiently large to constitute a particle reservoir for itself, the chemical potentials of the two components are also interrelated by

$$\mu_{\mathrm{M}_u\mathrm{O}_v}(T) = u\mu_{\mathrm{M}}(T) + v\mu_{\mathrm{O}}(p_{\mathrm{O}_2}, T), \tag{4.7}$$

where $\mu_{\mathrm{M}_u\mathrm{O}_v}(T)$ is the Gibbs energy per formula unit (f.u.) of the metal oxide. As a good approximation the weak pressure dependences of solid phases are not considered in Eq. 4.7 [178, 179]. Next, one has to consider that the decomposition reaction of the material into its most stable elemental constituents,

$$\mathrm{M}_u\mathrm{O}_v \leftrightharpoons u\mathrm{M} + \frac{v}{2}\mathrm{O}_2, \tag{4.8}$$

is always possible. Therefore, the chemical potentials are restricted to a certain range. For example, if the chemical potential of the oxygen is lowered by increasing the

temperature and decreasing the oxygen pressure (Eq.4.5), it is possible that via Eq.4.7 the metal chemical potential attains a value larger than the Gibbs energy per atom of the element ($\mu_{\mathrm{M}}(T) \geq \mu^{el}_{\mathrm{M}}(T)$). In this case, decomposition is favorable and all oxygen would be released from the oxide.

However, when the oxygen chemical potential is increased by increasing the partial pressure, it could be ultimately favorable to form an oxygen molecular liquid in coexistence with the metal. In that case, the chemical potential of the oxygen would exceed the the Gibbs energy per atom in an oxygen liquid ($\mu_{\mathrm{O}}(T) \geq \mu^{el,l}_{\mathrm{O_2}}(T)$). Such a limiting case is unrealistic and not obtained in relevant experiments. Thus, the decomposition at high oxygen chemical potential will not be observable in practice. As a consequence of this complication it is common to choose the cohesive energy of the oxygen molecule as an artificial but rigorous upper bound. This bound does not correspond to anything in nature, but such a chemical potential can be obtained in experiment via exposure to oxidizing gases like atomic oxygen or ozone at certain temperature and partial pressures. Fig.17 illustrates the dependence of oxygen chemical potential as a function of temperature at different oxygen partial pressures for dimeric oxygen and ozone. The data has been taken from JANAF electrochemical tables [180]. These chemical potentials at standard conditions (μ°_i of Eq. 4.5) were calculated using the most accurate quantum chemical methods and the pressure dependence is obtained via the ideal gas law (Eq. 4.5). Zero energy corresponds to the cohesive energy of the cold (0K) oxygen dimer. Using these gases, it is possible to cover the whole range of oxygen chemical potential, even those which are positive and exceed the bound previously defined.

In summary, the bounds for the chemical potentials are

$$\text{Oxygen-rich: } \mu_{\mathrm{O}} = \mu^{\mathrm{O_2}}_{\mathrm{O}}(T, isolated) \tag{4.9}$$

$$\text{Metal-rich: } \mu_{\mathrm{M}} = \mu^{\mathrm{M},el}_{\mathrm{M}}(T, p = 0). \tag{4.10}$$

In the zero Kelvin approximation these quantities can be calculated via electronic structure methods so that these states can serve as well defined reference states. Restating Eq. 4.7 using these references leads to

$$\begin{aligned} \mu_{\mathrm{M}_u\mathrm{O}_v}(T) &= u\left(\mu^{\mathrm{M},el}_{\mathrm{M}}(T=0) + \Delta\mu_{\mathrm{M}}(T)\right) \\ &+ v\left(\mu^{\mathrm{O_2}}_{\mathrm{O}}(T=0, isolated) + \Delta\mu_{\mathrm{O}}(p_{\mathrm{O_2}}, T)\right), \end{aligned} \tag{4.11}$$

where the Gibbs energy per f.u. of the compound ($\mu_{\mathrm{M}_u\mathrm{O}_v}(T)$) can be expanded in the following way

$$\begin{aligned} \mu_{\mathrm{M}_u\mathrm{O}_v}(T) &= H_{\mathrm{M}_u\mathrm{O}_v}(T=0) \\ &+ [H_{\mathrm{M}_u\mathrm{O}_v}(T) - H_{\mathrm{M}_u\mathrm{O}_v}(T=0)] \\ &- TS_{\mathrm{M}_u\mathrm{O}_v}(T). \end{aligned} \tag{4.12}$$

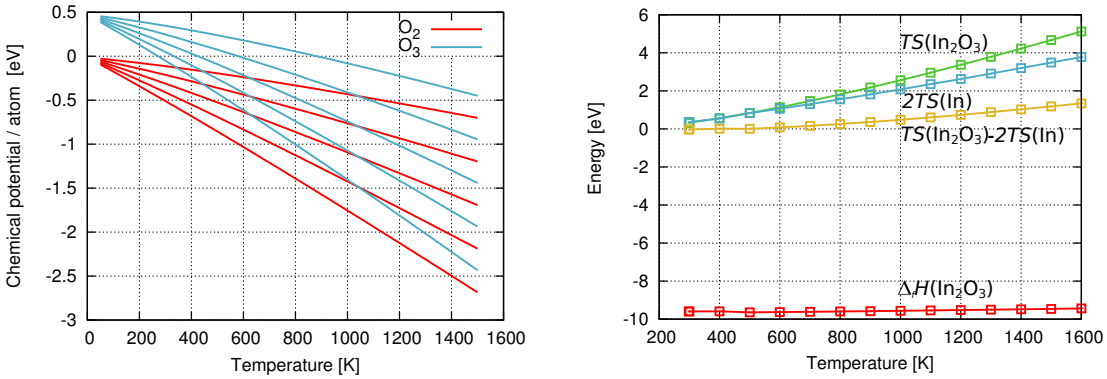


Figure 17: (left) The dependence of oxygen chemical potential as a function of temperature for partial pressures between 10^{-10}Pa and 10^{10}Pa. Steeper lines correspond to lower pressures. (right) Temperature dependence of the compound heat of formation and the entropy contributions to the free energies of In_2O_3 and metallic In. The difference of entropy contributions to the free energy between In_2O_3 and metallic In are also shown.

Introducing the compound heat of formation at zero temperature

$$\Delta_f H_{\mathrm{M}_u\mathrm{O}_v}(T=0) = H_{\mathrm{M}_u\mathrm{O}_v}(T=0) - u\mu_{\mathrm{O}}^{\mathrm{O}_2}(T, isolated) - v\mu_{\mathrm{M}}^{\mathrm{M},el}(T), \tag{4.13}$$

results together with Eq. 4.11 in

$$\begin{aligned} \Delta_f H_{\mathrm{M}_u\mathrm{O}_v}(T=0) \quad &+ \quad [H_{\mathrm{M}_u\mathrm{O}_v}(T) - H_{\mathrm{M}_u\mathrm{O}_v}(T=0)] - TS_{\mathrm{M}_u\mathrm{O}_v}(T) \\ &= \quad u\Delta\mu_{\mathrm{M}}(T) + v\Delta\mu_{\mathrm{O}}(p_{\mathrm{O}_2}, T). \end{aligned} \tag{4.14}$$

Apart from zero-point vibrations and the zero pressure condition, Eq. 4.14 is exact and gives a relation between the chemical potential of the constituents and thermochemical data which can be obtained from calorimetric data. The expression can be simplified by recognizing that the temperature dependence of the compound heat of formation is insignificant,

$$\Delta_f H_{\mathrm{M}_u\mathrm{O}_v}(T=0) \quad - \quad TS_{\mathrm{M}_u\mathrm{O}_v}(T) \tag{4.15}$$

$$= \quad u\Delta\mu_{\mathrm{M}}(T) + v\Delta\mu_{\mathrm{O}}(p_{\mathrm{O}_2}, T), \tag{4.16}$$

as it is also illustrated in Fig. 17 for the case of In_2O_3. In contrast, the entropy term of the compound is more significant (Fig. 17). The (vibrational) entropy of the metallic constituent can be split from $\Delta\mu_{\mathrm{M}}(T)$ and grouped together with the entropy of the compound:

$$\begin{aligned} \Delta_f H_{\mathrm{M}_u\mathrm{O}_v}(T=0) \quad &- \quad T\left[S_{\mathrm{M}_u\mathrm{O}_v}(T) - uS_{\mathrm{M}}(T)\right] \\ &= \quad u\Delta\mu_{\mathrm{M}}(T=0) + v\Delta\mu_{\mathrm{O}}(p_{\mathrm{O}_2}, T). \end{aligned} \tag{4.17}$$

The term $T\left[S_{M_uO_v}(T) - uS_M(T)\right]$ is plotted in Fig. 17 and demonstrates that the entropies of the solids largely tend to cancel. The resulting approximative relation between the so called thermodynamic connection energies and the (calculated) compound heat of formation at $T = 0$K is

$$\Delta_f H_{M_uO_v}(T=0) = u\Delta\mu_M(T=0) + v\Delta\mu_O(p_{O_2}, T). \tag{4.18}$$

The error in this relation at higher temperatures is illustrated in Fig. 17 for the case of In_2O_3. Note that the deviations can be as large as 1 eV at temperatures of 1500K. The deviations at lower temperatures are usually smaller compared to uncertainties connected with the DFT approximations. The reason for this is that the calculation a quantity like the heats of formation involve energy differences between materials with different types of bonding (metal, molecule and semiconductor). In these cases error cancellations, which are very important for the success of DFT based methods, are not pronounced. Especially, in the case of molecules the total energy calculations are often problematic. First, molecules are no periodic objects and have a localized electron density in space. Furthermore, open shell systems like oxygen are generally less reliable, because in these systems exchange and correlation effects become more important.

Finally, the bounds determined by Eq. 4.18 are not attainable in practice, if other stable oxides exist with different stoichiometry. In those cases additional constraints apply for the chemical potentials in the form

$$\Delta_f H_{M_nO_m}(T=0) > n\Delta\mu_M(T) + m\Delta\mu_O(p_{O_2}, T). \tag{4.19}$$

4.2.2 *Electron reservoir*

According to Eq. 4.4 the formation energy of point defects also depends on the chemical potential of the electrons. This is because charges can be transfered between the defect site and the continuum of the bulk states. In semiconductor physics and in the absence of electric fields, the chemical potential of the electrons is often called the Fermi-energy. Equation 4.4 requires an absolute definition of the Fermi energy which is different from semiconductor physics where this level is usually referenced with respect to one of the band edges. In order to enable referencing with respect to *e.g.* the VBM of the given material the electron removal or hole creation energy is needed. This energy corresponds to the ionization potential (I_p) of the material, which is not defined for strictly periodic systems. Nevertheless, in DFT the ionization potential can be calculated with respect to the internal reference using total energy differences

$$I_p^{DFT} \equiv E_{VBM} = \lim_{N\to\infty} E(N-1) - E(N) \tag{4.20}$$

It was shown that for the highest occupied level (VBM), the orbital energy itself can be identified with I_p^{DFT}[181, 182, 183, 184, 185]. This is known as the DFT version

of Koopman's theorem, which applies in Hartree-Fock theory for any orbital [186, 187, 147] . Nevertheless, the result is remarkable, since the eigenvalues of the Kohn-Sham equations (Eq. 3.8) were simply introduced as Lagrange multipliers to ensure orthonormality. Thus, it is justified to attach limited physical meaning to the orbital energies. The I_p referenced with respect to the vacuum level can be obtained via

$$I_p^{vac} = I_p^{DFT} - E_{vac}^{DFT}, \tag{4.21}$$

where E_{vac}^{DFT} is the vacuum electrostatic potential within a DFT surface calculation. The I_p^{vac} can directly be compared with experiments and contains the dipole contributions of the surface used for the calculation.

Fig. 18 shows the electron removal energy for germania calculated in a finite supercell using 72 atoms for a fractional charge ranging between $q = 1$ and $q = \frac{1}{10000}$*. It can be seen that the energy converges to the highest Kohn-Sham orbital eigenvalue, which is denoted by black horizontal lines.

Although it was shown that at least the VBM of a DFT calculation is meaningful, the band gaps are usually severely underestimated. This deficiency is the band gap problem of DFT [188, 189] and well illustrated with GeO_2 in Fig. 18. The DFT-LDA band gap is 1.8 eV compared to 5.6 eV [92] in experiment. It can be argued that this is because Koopman's theorem does not apply for the lowest unoccupied orbital. This is, however, only partially true, because the electron affinity can also be obtained using total energy differences showing a similar underestimation. The reasons for the underestimation of band gaps ($\geq$ 50% in LDA/GGA) are rooted in the crude approximations used in the exchange correlation functionals, *i.e.* the spacial locality of the LDA or GGA functionals. A large portion of the band gap error arises due to the so-called (functional) derivative discontinuity of the XC-energy [188, 189, 190] at integer electron numbers,

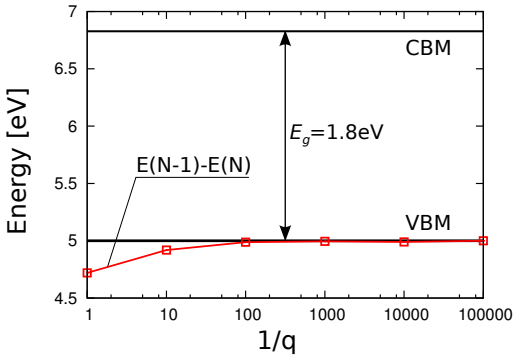


Figure 18: The fractional electron addition and removal energies for GeO_2.

$$\delta E = \frac{\delta E_{XC}[n(r)]}{\delta n(r)}. \tag{4.22}$$

In contrast to the real behavior, the local XC-functionals provide a smooth convex behavior of the energy as a functional of electron density. It is, however, encouraging that by using more complex forms for the XC-potential (*e.g.* hybrid functionals) the discontinuous behavior can be restored [191].

*This procedure is equivalent to increasing the system size and removing one electron.

4.3 FINITE SIZE EFFECTS

As already discussed in Sect. 3.3 the plane wave approach greatly simplifies the calculations by using periodic boundary conditions. At the same time some unphysical energy contributions are introduced by this method in defect calculations [192], which have to be corrected. When defects are constructed within the supercell, they are present in all image cells, giving an average defect concentration of about $10^{-21}\text{cm}^{-3} - 10^{-22}\text{cm}^{-3}$. These values are usually not realistic for intrinsic point defects. There are basically four different effects which change the total energy of such a calculation. If a charged defect is introduced, there is an electrostatic interaction between the periodic images and the so called jellium compensation charge. At a sufficiently high distance it can be expressed by a multipole expansion [193]. The dominating term is the monopole interaction, which for cubic crystals can be written in terms of a Madelung energy (first term in Eq. 4.23):

$$\Delta E(q, \epsilon_0) = +\frac{q^2 \alpha_M}{2\epsilon_o V_c^{1/3}} + \frac{2\pi q Q}{3\epsilon_0 V_c} + \mathcal{O}(V^{-5/3}). \tag{4.23}$$

Q is the second radial moment of the defect charge density within the super cell with the volume V_c and ϵ_r is the static dielectric constant. In general the first order contribution is represented by an Ewald summation consisting of a periodic point charge compensated by a so called homogeneous jellium background in order to avoid infinite energy contributions. The following terms in Eq. 4.23 are higher multipoles which generally decay as $\mathcal{O}(V^{-1})$ or faster compared to the $\mathcal{O}(V^{-1/3})$ for monopole corrections. The effect of the higher terms is opposite to that of the monopole term and typically $\sim 30\%$ in magnitude.

The jellium compensation charge is applied in practice by setting the constant offset of the electrostatic potential expansion to zero $(V(\mathbf{G} = 0))$. This procedure leads to a detuning of the energy spectrum as well as the electrochemical potential and has to be realigned. The effect as such does not influence the total energy of the system, because the offset in the energy spectrum is always counterbalanced by the double counting terms in the evaluation of the total energy [192, 147]. In the case of charged defect calculations, however, the E_{VBM} enters as an extra term into the calculation of the formation energies (Eq. 4.4) because it is the reference energy for the Fermi level.

Besides electrostatic interactions, there are also band structure effects introduced by the defect. Because of the high concentration, there is the possibility of a defect band with a significant dispersion. Of course, this dispersion would not occur in reality, because it requires an overlap of defect state wave functions. Therefore, in principle higher occupancies of defect states have to be subtracted from the total energy in order to obtain the energy value of the dilute limit.

Several methods were proposed in order to deal with the various finite-size effects of point defects [193, 194, 192, 146, 195, 196, 197, 198, 199]. Most of the methods aim at obtaining the correct energies with already small supercell sizes in order to reduce

computational cost. Therefore, analytic expressions such as *e.g.* the Madelung energy of a point charge (resembling the defect charge) embedded in a jellium are added to the defect formation energies [193, 194, 192]. The reliability of such methods cannot be guaranteed for any defect.

Only recently a method has been proposed to calculate a correction for the electrostatic interactions and the potential disalignment in an ab initio fashion using the electrostatic potential and charge density from the calculations [198]. This is probably the most elegant way to deal with the problem. For the present study, however, this method was not available early enough so that the more tedious but most reliable method has been adopted. In the extrapolation scheme defects are calculated in increasingly larger supercells until an analytic scaling behavior is observed. In the case of charged supercells this is usually $E \propto V^{-1}$. Once this scaling behavior is visible the formation energy is extrapolated to infinity. A very thorough evaluation of this approach was recently carried out for ZnO [200, 196].

4.4 VIBRATIONAL PROPERTIES

The phonon density of states can be obtained via electronic structure methods using different approaches. At the heart of all methods are the Born-Oppenheimer approximation [61] and the harmonic approximation of the potential energy (U) as a function of small displacements from the equilibrium crystal lattice positions $\mathbf{u}(\mathbf{R})^{\dagger}$,

$$U_{\text{tot}} = U_{\text{eq}} + U_{\text{harm}}(\mathbf{u}\{\mathbf{R_I}\}). \tag{4.24}$$

The harmonic term is a function of the displacements from equilibrium positions and can be written as

$$U_{\text{harm}}(\mathbf{u}\{\mathbf{R}\}) = \frac{1}{2}\sum_{i,j,\nu,\mu}\frac{\partial^2 U_{\text{tot}}}{\partial u_{i\nu}\partial u_{j\mu}}u_{i,\nu}u_{j,\mu} = \frac{1}{2}\sum_{i,j,\nu,\mu}\mathbf{u}_{i\nu}\mathbf{C}_{i\nu,j\mu}\mathbf{u}_{j\mu}. \tag{4.25}$$

In this equation $\mathbf{C}_{i\nu,j\mu}$ is the so called force constant matrix (FCM) and contains $3N_{\text{at}} \times 3N_{\text{at}}$ elements, where N_{at} is the number of atoms in the system. The i, j are the atomic indices whereas ν, μ correspond to different cartesian directions (x, y, z). Once the FCM is known, for each atom and cartesian direction one equation of motion can be formulated

$$M\ddot{\mathbf{u}}_{j\mu} = -\sum_{i,\nu}\mathbf{C}_{i\nu,j\mu}\mathbf{u}_{i\nu}. \tag{4.26}$$

The solution to to this equation can be written in the form of plane waves

$$\mathbf{u}\,(\mathbf{R}, t) = \epsilon e^{i(\mathbf{k}\cdot\mathbf{R}-\omega t)}, \tag{4.27}$$

[†]Bold letters refer to vector and matrix notation

where ϵ is the polarization vector and describes the direction in which the ions move for a particular vibrational mode and $\mathbf{k}$ and ω are the wave vector and frequency, respectively. Inserting Eq.4.27 into Eq.4.26 results in the eigenvalue equation

$$M\omega^2\epsilon = \mathbf{D}(\mathbf{k})\epsilon, \tag{4.28}$$

with the Fourier transform of the FCM called the dynamical matrix

$$\mathbf{D}(\mathbf{k}) = \sum_{\mathbf{R}} \mathbf{C}(\mathbf{R}) e^{i(\mathbf{k}\cdot\mathbf{R}-\omega t)}. \tag{4.29}$$

For a given wave vector it is thus possible to obtain the frequency spectrum of the system with $3N$ eigenvalues. The frequencies calculated at a limited set of $\mathbf{k}$-vectors can be interpolated to construct the phonon dispersion relation $\omega_n(\mathbf{k})$ for the n different phonon branches. The phonon density of states as a function of frequency is obtained as the histogram over the branches at a sufficiently large set of $\mathbf{k}$-points

$$g(\omega)\frac{1}{3N_{\text{at}}N_{\text{norm}}}\sum_{n,\mathbf{k}}\delta(\omega-\omega_n(\mathbf{k})). \tag{4.30}$$

The $\mathbf{k}$-points are usually selected using the Monkhorst-Pack scheme [201] whereas the normalization factor N_{norm} is chosen to satisfy the condition

$$\int_0^\infty g(\omega)d\omega = 1. \tag{4.31}$$

All phonon related thermodynamic properties can be obtained from $g(\omega)$ when combined with the Bose-Einstein occupation function. In particular the phonon entropy of the system can be expressed as [179]

$$\begin{aligned} S^{vib} &= N_f k_B T \int_0^\infty \left[\frac{\hbar\omega}{2k_BT}\coth\left(\frac{\hbar\omega}{2k_BT}\right)\right. \\ &- \left.\ln\left(2\sinh\left(\frac{\hbar\omega}{2k_BT}\right)\right)\right] g(\omega)d\omega, \end{aligned} \tag{4.32}$$

with N_f being the number of degrees of freedom in the calculated system.

The main ingredient in the formalism is the FCM or its analogue in Fourier space, the dynamical matrix. Two approaches exist in order to calculate the entries of these matrices. The most simple approach is the frozen phonon method also used in the present thesis. A series of calculations is performed where different atoms are slightly displaced from their equilibrium positions. For each given displacement, the forces are calculated on all atoms in the supercell using the Hellmann-Feynman theorem [161]. Together with the displacement length, the forces can be converted into force constants. It can be noted that the Hellman-Feynman theorem itself is a result of

first order perturbation theory, for which knowledge of the ground state wave function is sufficient. The second method is a generalization in this respect,because the second derivatives (force constants) are directly calculated using second order perturbation theory. This, however, can only be obtained using the first order wave functions of the system. One of the greatest advantages of this density functional perturbation theory [202, 203, 204] (DFPT) as compared to other non perturbative methods for calculating the vibrational properties of crystalline solids, is that within DFPT the responses to perturbations of different wave lengths are decoupled. This is possible because the second order response is directly evaluated in Fourier space *i.e.* the entries of the dynamical matrix. This feature allows one to calculate phonon frequencies at arbitrary wave vectors avoiding the use of supercells and with a workload that is essentially independent of the phonon wave length.

The reason for choosing the frozen-phonon method in the present study is rooted in the fact that vibrational properties of defects are calculated. In that case, supercells have to be used in order to avoid finite-size errors due to the defect so that there are no advantages in using DFPT.

4.5 POINT DEFECT KINETICS

4.5.1 *Self-diffusion constant*

In order for the material to reach equilibrium, defects have to migrate to natural sinks/sources within the material. In practice this may or may not be possible depending on the defect kinetics, annealing time and temperature. If defects occupy local minima in the total energy landscape, transition state theory [205, 206, 207] is applicable so that the jump frequency of an atom can be expressed in the simple form

$$\Gamma = \Gamma_0 \exp\left(-\frac{\Delta G^m}{k_B T}\right) = \Gamma_0 \exp\left(\frac{\Delta S^m}{k_B}\right) \exp\left(-\frac{\Delta H^m}{k_B T}\right), \tag{4.33}$$

where Γ_0 is the so called attempt frequency and ΔG^m, ΔS^m and ΔH^m are the free energy, entropy and enthalpy of migration. Further, in the small vibration approximation this can further be decomposed into

$$\Gamma_* = \Gamma_0 \exp\left(\frac{\Delta S^m}{k_B}\right) = \left(\frac{\prod_{(j=1)}^{N} \nu_j^{eq}}{\prod_{(j=1)}^{N-1} \nu_j^{TST}}\right). \tag{4.34}$$

The effective frequency Γ_* is very different from a simple Einstein frequency [61], or any single frequency exhibited in physical space. It is the ratio of the product of the N normal frequencies of the entire system at the starting point of the transition

to the $N-1$ normal frequencies ν_j of the system constrained in the saddle point configuration [205]. Due to Eq. 4.33, which is valid at microscopic level, also the diffusion constant has a similarly simple form

$$D^d = D_0^d \exp\left(-\frac{\Delta G_{eff}^m}{k_B T}\right). \tag{4.35}$$

In simple cases a straightforward connection can be made with the microscopic theory to express $D_{0,eff}$ and ΔG_{eff}^m as functions of transition frequencies of individual atomic processes. In the case that movements of defects are not correlated, the diffusion constant of one defect d in a certain crystal direction μ can be expressed as [207]

$$D_\mu^d = \frac{1}{2} \sum_i Z_i^d \, \Gamma_i^d |\lambda_{i,\mu}|^2, \tag{4.36}$$

where the sum runs over all microscopic processes, Z_i^d is the multiplicity of the process, Γ_i^d are the frequencies from Eq. 4.33 and $|\lambda_{i\mu}|$ is the projected length of the process on the direction μ. Due to the exponential term, the diffusivities are largely determined by the lowest barriers and their corresponding diffusion parameters. In the case of more complicated structures, where migration processes are correlated to a certain degree, the evaluation is more difficult. As a general relation it is, however, always possible to employ the Einstein relation

$$D_\mu = \lim_{t\to\infty} \frac{\langle R_\mu^2 \rangle}{2t}, \tag{4.37}$$

where the t is the diffusion time and R_μ^2 is the projected mean square displacement of the defect. According to the frequencies determined by Eq. 4.33 it is possible to perform a random walk in a given structure. This can be done by solving the Master-Equation [207] for this problem or alternatively using the Kinetic Monte-Carlo algorithm [208]. The obtained diffusion constant as a function of temperature can thus be used to determine the effective migration parameters for Eq. 4.35. The major ingredient necessary from electronic structure calculations are the migration enthalpies[‡] which can be determined using methods like the nudged elastic band (NEB) or the dimer method.

4.5.2 *Nudged elastic band method*

The Nudged Elastic Band (NEB) [209, 210, 211] method is used to identify reaction pathways when both the initial and final states are known. As a first guess, several intermediate steps (images) between the initial and the final state are created, where linear interpolation is often a sufficient approximation. Then all images

[‡]The zero pressure condition is again appropriate in the case of migration enthalpies

are relaxed with "spring forces" acting in between. This leads to a constrained structural relaxation in $3N-1$ dimensions, where N is the number of particles in the simulation box. The constrained direction is the so called tangent to the image chain, which is along the low energy path once the method has converged. Further, with the climbing image modification to this basic algorithm it is possible to calculate not only the low energy path for the reaction but also the transition state (TS) configuration at the saddle point. The climbing image (CI-NEB) method uses inverted forces along the tangent to the image chain for the image closest to the saddle point. This drives the image with the highest energy exactly to the saddle point during structural optimization.

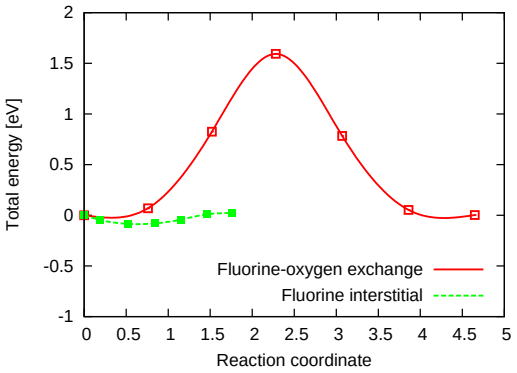


Figure 19: Calculated results for fluorine migration in a SnO_2 host lattice with 72 atoms.

Fig.19 shows the calculated total energies of five images along two different reactions path as an example. They correspond to diffusion processes of fluorine in SnO_2 [108]. The larger barrier corresponds to oxygen vacancy mediated diffusion of substitutional fluorine, whereas the shorter path corresponds to fluorine interstitial diffusion. In the latter case, the barrier is slightly negative suggesting a different ground state from that assumed prior to the barrier calculation. Because the CI-NEB method was used for the calculation, it is justified to assume that the image with the highest energy is the saddle point configuration, even if the image density along the path is low.

4.6 SURFACES

4.6.1 *Slab calculations*

Surface calculations are commonly conducted within the supercell approach. In order to obtain free surfaces within periodic boundary conditions a vacuum region is inserted into the supercell. The typical geometry is schematically depicted in Fig. 20 with the typical dimensions employed in practical calculations. As in the case of defects, there can be finite-size errors associated with this *slab geometry* due to interactions of the surfaces through the material and through the vacuum. The interactions are especially strong for polar surfaces (Sect. 4.6.3). Generally, the interactions via the vacuum decay sufficiently fast. The interaction through the slab is only strong for metallic systems due to the confinement effect on delocalized states.

Fig. 21 illustrates the convergence behavior of the energy with respect to these two parameters in the case of In_2O_3-(001) surfaces. Reliable energies can be extracted from relatively thin slabs ($\sim$ 10Å). The fast convergence of the energy with respect to the vacuum width is surprising, because of the expected long ranged $(1/r)$ behavior of the electrostatic potential at real surfaces of materials [149]. In contrast, local functionals like the LDA or GGA produce an exponential decay of the electrostatic potential which is unphysical, but beneficial for surface calculations.

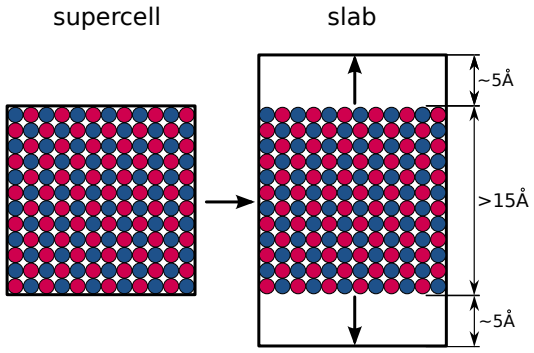


Figure 20: Typical slab geometry and size in the case of non-metallic non-polar materials.

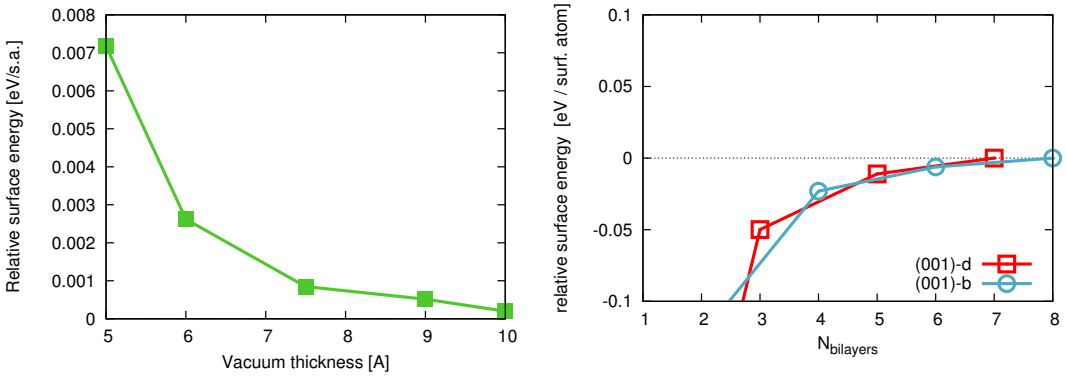


Figure 21: (left) Convergence of the slab energy as a function of vacuum width. (right) Convergence of the energy as a function of the slab thickness for two different In_2O_3 terminations.

4.6.2 *Surface thermodynamics*

The thermodynamics of surfaces is analogous to that of point defects as described in Sect. 4. Also for surfaces a grand canonical ensemble is adopted in order to compare surfaces with varying stoichiometry. In contrast to defect calculations, charging of the surface is usually not considered, although it might be important on real surfaces. This is, because a hypothetically unsaturated surface can take up a large amount of charge leading to strong finite size effects.

The surface tension is defined as the excess of the grand potential per unit area per surface

$$\gamma_G = \frac{1}{A}\left\{\Phi_{G,\text{slab}} - \Phi_{G,\text{ideal}}\right\} \tag{4.38}$$

$$= \frac{1}{A}\left\{E_{\text{slab}} - E_{\text{ideal}} - TS_{\text{slab}} - S_{\text{ideal}} - \sum_i \mu_i(n_{\text{slab}} - n_{i,\text{ideal}})\right\} \tag{4.39}$$

$$= \frac{1}{A}\left\{\Delta E_{\text{surf}} - T\Delta S_{\text{surf}} - \sum_i \mu_i \Delta n_{i,\text{surf}}\right\} \tag{4.40}$$

The isochoric version of the grand potential is more appropriate in the case of surfaces since it is assumed that the surface is supported by a sufficiently large portion of the bulk and lateral relaxation is not possible. Like in the case of point defects the surface energy ΔE_{surf}, and the surface entropy ΔE_{surf}, can be calculated using electronic structure methods. In the present thesis, however, the entropy is not evaluated. This is justified, because the absolute values of surface tension are only of limited interest. More important are the relative values for different terminations as well as the relative surface tensions between different stoichiometries of the same material. Thus, while in point defect calculations the entropies of the ideal material are compared with the defective material, in the case of surface calculations different defected structures are compared among each other leading to smaller entropy differences [212].

4.6.3 *Polarity*

Generally, surfaces of ionic materials can be categorized according to the stacking sequence along the surface normal. This classification was established first by Tasker [213, 214, 215] who divided surfaces into three different classes. In this scheme crystal planes are gathered into charge neutral groups. In the most simple case each individual atomic plane is effectively charge neutral resulting in a stacking of neutral units and referred to as type-I surfaces.

For type-II surfaces the charge neutral unit consist of at least a symmetric trilayer where each plane contains excess charge. The important point concerning type-II surfaces is that the charge neutral group of planes contains no net dipole.

Finally, for some crystallographic directions it is not possible to group planes charge neutrally without a dipole. Such surfaces are referred to as polar. For type III surfaces a subdivision can be made. In some cases it is possible to divide the planes of one charge in half planes and construct dipole free trilayers using only half planes of one charge type and full planes of the opposite charge. In a slab calculation this is done by transferring half of the surface atoms from one side to the other side of the material as it is depicted in Fig. 22. This surface shall be referred to as type-IIIa in subsequent chapters. Crystal structures of such materials contain an inversion symmetry or some mirror plane perpendicular to the surface normal and are not strictly polar materials

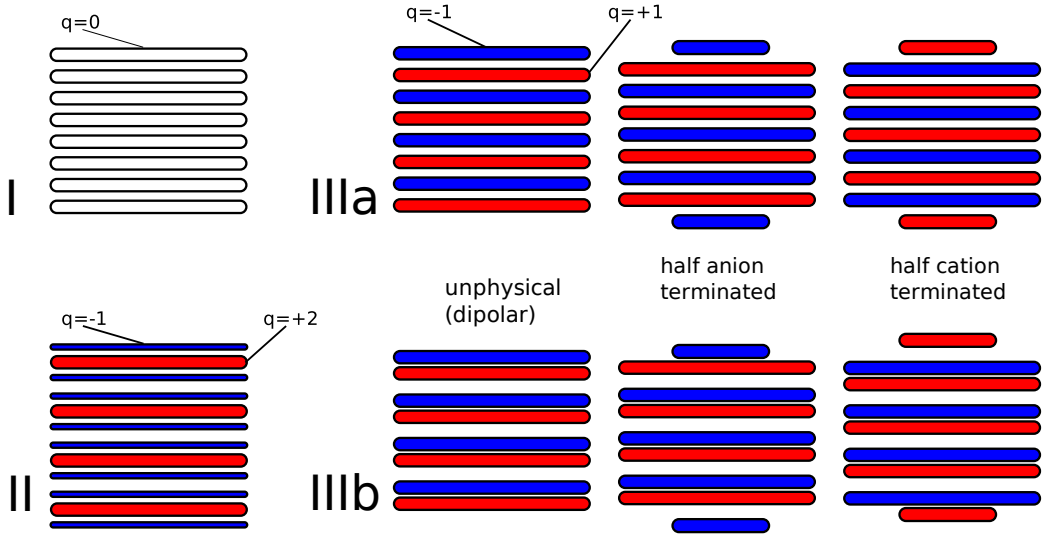


Figure 22: Surfaces classified according to Taskers scheme of ionic compounds. In addition to the original classification the third type has been subdivided in two subclasses.

in a sense that they would exhibit a piezoelectric response. Nevertheless, such surfaces exhibit characteristics similar to those lacking these symmetry elements. In truly polar cases (type-IIIb), the dipole cannot be removed from the slab by any redistribution. The different types of surfaces are depicted in Fig. 22 with their essential features.

From an electrostatic point of view it is assumed that the surface energy for the different classes of surfaces is ordered as $\gamma(I) < \gamma(II) < \gamma(IIIa) < \gamma(IIIb)$. It can be shown that for the type-III surfaces the energy diverges as a function of the materials thickness. In the case of type-IIIa surfaces this divergence is removed by the rearrangement of the atoms. Clearly, the divergence exists only in a purely ionic picture, where no polarization, relaxation or other processes are allowed. In a real semiconductor or insulator at a certain energy density (thickness) the threshold for some compensation mechanism will be reached such as the creation of defects, large structural relaxations, reconstructions or band gap excitations. Nevertheless, the associated surface energies of these surfaces are assumed to be high compared to non-polar surfaces. Finally, it should be remarked that Tasker's model of charged planes is equivalent to the electron counting model [216, 217, 218, 219] which is more popular in the case of more covalent compound semiconductor surfaces.

Part III

POINT DEFECT THERMODYNAMICS

The following part of the thesis deals with the thermodynamic stability of point defects in In_2O_3 and SnO_2. Starting with an approximation with additional semi-empirical on-site corrections for exchange and correlation we explore the stability of a wide range of defect arrangements. We also estimate the magnitude of entropy contributions to the defect formation energy for selected defects. In the second chapter, we then focus on the influence of the XC-functional in defect calculations and employ a non-local form which is computationally more demanding but adds some missing physics to the calculations. Namely, it is the band gap error which is corrected by this method in a physically motivated manner. As a test case we apply different XC-approximations to the oxygen vacancies of SnO_2, In_2O_3 but also ZnO as a benchmark. Finally, the new methodology is used to obtain the formation energies of point defects in SnO_2 which enables us to compare the defect equilibria in these two materials. We conclude by using the refined formation energies in order to obtain maximum doping levels for the two materials.

5

INTRINSIC POINT DEFECTS OF In_2O_3: A CASE STUDY WITHIN GGA+U

5.1 INTRODUCTION

In this first chapter a survey of possible point defects is presented. This study serves as a guide for the upcoming chapters where the results are further refined. Since In_2O_3 is the prototype TCO material most frequently used up to now [3] it serves as a reference for comparisons with SnO_2 and other TCOs.

Although previous first principles studies on point defects of In_2O_3 exist [128, 129, 130, 131] the present treatise is the most comprehensive study up to now. Furthermore, a modified version of the GGA-PBE functional is employed: the GGA-PBE+U [220, 168, 167]. The computational setup of this method was previously tested in a band structure study [221]. Further details on the computational setup and numeric accuracy as well as a comparison with experimental material properties can be found in Ref. [A5]. Using this method, an improved agreement for the valence electronic density of states is obtained with measured photoelectron spectroscopic data [222, 223].

A second difference to previous investigations is that not only the defect formation energies, but also formation entropies are presented for several defects. To the author's knowledge it is the first attempt to calculate formation entropies within a plane wave based supercell approach. Previous studies relied on computationally less demanding methods like atomistic potentials [224, 225, 226, 227, 228, 229] or tight binding approaches [230, 231, 232].

In the case of formation energies, we have included a large variety of defect types starting with the cation and anion vacancies as well as the interstitials. In addition, we have investigated the anti-site defects as well as several arrangements of defect pairs and dumbbell geometries for cation and anion interstitials. The calculations encompass a wide range of charge states for all defect geometries.

Finite size effects were corrected using methods described in Ref. [A5] and Refs. [146, 193, 195, 192] with supercells containing 40, 80, 120 and 160 atoms.

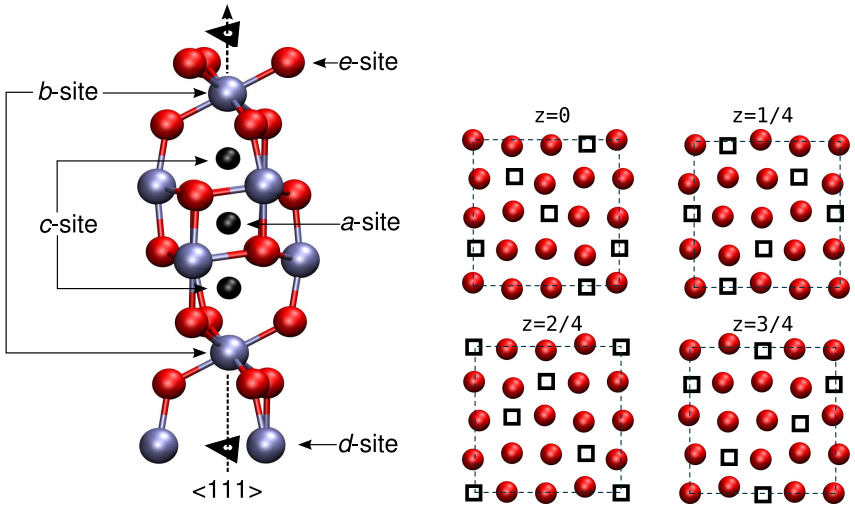


Figure 23: (left) Arrangement of atoms surrounding the ⟨111⟩ threefold axis. Black spheres denote interstitial sites enclosed by two regular indium *b*-sites in the order *b*-*c*-*a*-*c*-*b*. Indium atoms are also located on *d*-sites which have a two-fold point symmetry. Oxygen atoms reside on the low-symmetry *e*-site. (right) Distribution of structural vacancies within the oxygen (001) layers.

5.2 CRYSTAL STRUCTURE

Under ambient conditions In_2O_3 assumes the bixbyite structure ($Ia\bar{3}$, space group no. 206) [68]. The primitive unit cell contains 40 atoms whereas the conventional (bcc) unit cell contains 80 atoms. Important structural elements of this lattice are depicted in Fig. 23. Namely, the atomic arrangement around the ⟨111⟩ threefold rotation axis is shown along with the labeling of different sites. On the right, the arrangement of the interstitial *c*-sites within (001) planes are shown, a structural element usually termed as "structural vacancies". This is because in the related CaF_2 structure these sites are occupied by anions. Besides the *c*-site, interstitials can be placed on *a*-sites. In addition, split interstitials are possible along the threefold and twofold rotation axes on In-*b* and In-*d* sites, respectively.

5.3 FORMATION ENERGIES

The Fermi level dependence of the formation energies calculated using Eq. 4.4 (Chapter 4) is shown in Fig. 24 for both oxygen and indium-rich conditions. In the indium rich regime (reducing conditions) oxygen vacancies (V_O) display the highest stability. Their formation energy is negative throughout major parts of the band gap pinning the Fermi level above the gap center. Competing donor defects like indium interstitials (In_i) and indium anti-sites (In_O) do always have significantly higher formation energies

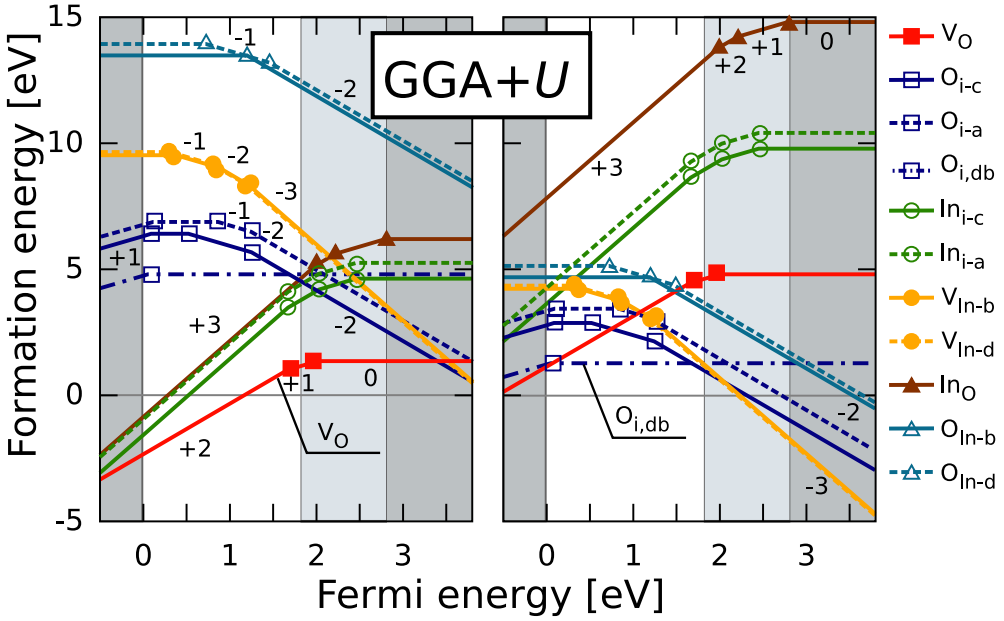


Figure 24: Fermi level dependence of formation energies under indium and oxygen-rich conditions calculated with GGA+U. The data shown include monopole and finite size corrections. The white strip indicates the calculated band gap 1.81 eV (GGA+U). The difference to the experimental band gap is shown by light grey shading. Dark areas represent the valence and conduction bands. The VBM is located at $E_f = 0$.

when compared with the oxygen vacancy. Higher formation energies are also found for all acceptor defects. The principal acceptor defect in In_2O_3 is the oxygen interstitial (O_i^{2-}), which is in line with the defect model of Frank and Köstlin [109, 112, 105, 233] (see Sect. 2.2.4) and previous GGA calculations [131]. The lowest formation energy for interstitials is obtained when placed on *c*-sites (Fig. 23).

An interesting finding is that for the highest Fermi energies possible, formation energies of O_i^{2-} and indium vacancies (V_{In}^{3-}) are about equal. This means that at the doping limit there are in fact two different acceptors compensating the extrinsic dopant. Notably, the V_{In}^{3-} and O_i^{2-} acceptor defects cross the zero energy line for Fermi levels located within the conduction band. Under oxygen-rich conditions (oxidizing conditions), neutral oxygen interstitials in a dumbbell conformation ($O_{i,db}$) are the dominant defects. Finally, the oxygen anti-sites always have formation energies higher than any other acceptor so that they can be excluded on the basis of these calculations.

5.4 GEOMETRY AND ELECTRONIC STRUCTURE

5.4.1 *Donor defects*

The defect relaxations around the $In_{i\text{-}c}$ tend to be inward for the highest and dominant charge state $q = +3$. The anions are pulled inward (by up to 16%) * while cations are pushed slightly outwards (2.5%). In less positive charge states, the relaxation tends to be outward. For V_O, the geometry hardly changes in the neutral charge state and a slight inward relaxation is found for the nearest indium ($\sim -2.5\%$). For higher charge states these cations are pushed outwards (up to 10%).

The electronic structure of V_O and In_i (Fig. 25 right) show some similarity. Both defects create a notable change in the band structure as compared with the ideal structure (Fig. 10). Apart from the fundamental band gap, which is decreased due to the introduction of the defects, a second gap emerges inside the conduction band. A similar effect has been observed for Sn dopants [234]. This additional gap is produced by the interaction of In-5*s* states with defect states. The gap is larger for V_O (1.47 eV for V_O compared with 1.03 eV for the In_i in neutral charge states). Figure 25 also shows the similar charge dependent level relaxation for the two donors. The band structures were calculated using a 40 atomic unit cell so that the band dispersion of the defect states is caused by overlap via the periodic boundary conditions. It is notable that the defect states are relatively dispersed and have a strong CB character.

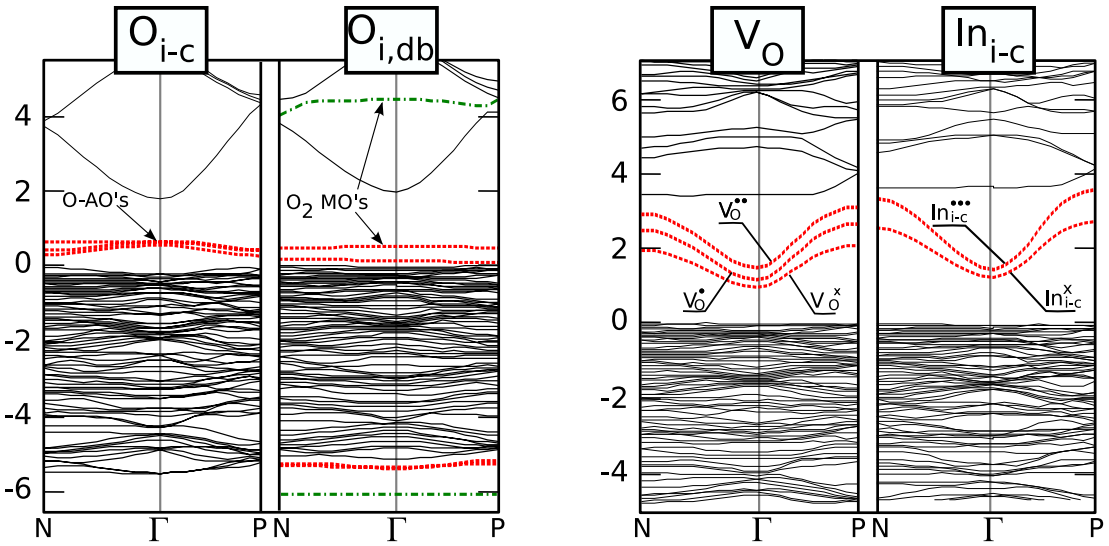


Figure 25: (left) Electronic band structures of oxygen interstitials on a *c*-site and in dumbbell configuration. (right) Band structures are shown with oxygen vacancy and indium interstitial in different charge states.

*Values are given with respect to original bond length.

5.4.2 *Acceptor defects*

Oxygen interstitial

The introduction of a neutral oxygen on interstitial sites leads to some outward relaxation (8 %). For negative charge states, due to strong charge localization, indium atoms move towards the interstitial site (8%) and oxygen neighbors are pushed farther outwards by 5%. It is interesting to note that O_i can exist in neutral and even positive form. In the neutral case, however, it is not favorable to localize an electron hole on the interstitial itself. Instead, neutral oxygen readily reacts with the oxygen sublattice in order to form a dumbbell configuration. Similar geometries were found previously for ZnO [235]. In this configuration the charge deficiency (hole) is evenly distributed on two oxygen atoms by the formation of a covalent bond, whereas other atoms in the structure remain largely unaffected. The thus obtained structure can be thought of as a peroxo-ion (O_2^{2-}). There are also various ways to accommodate the oxygen dumbbell within the host lattice as it is shown in Fig. 27. The orientation of the most stable dumbbell ($O_{i,db}$) lies within a {100} plane and is aligned (as much as possible) with two neighboring interstitial *c*-sites along $\langle 110 \rangle$ directions. The high electron concentration between the participating atoms indicates covalent bonding (Fig. 26).

In the electronic band structure two highly localized defect states for the dumbbell configuration (Fig. 25 left) can be identified, whereas the *c*-site regular interstitials exhibit three atomic orbital-like oxygen acceptor defect bands with strong *p*-character. In the first case the discrete energy levels of a peroxo molecule can be assigned, whereas in the latter case the states resemble the three atomic states of the isolated oxygen ion. The uppermost occupied defect bands are identified as the anti-bonding π^*_{p-p} molecular orbitals. The remaining oxygen dimer molecular orbitals are also highlighted in the figure. For this defect, adding electrons would result in the occupancy of the conduction band since the σ^*_{p-p} states can be found only 2.2 eV above the CBM. Therefore, no ambipolar behavior can be established for this defect type and the negative charge state of the dimer is artificial (Fig. 27). The

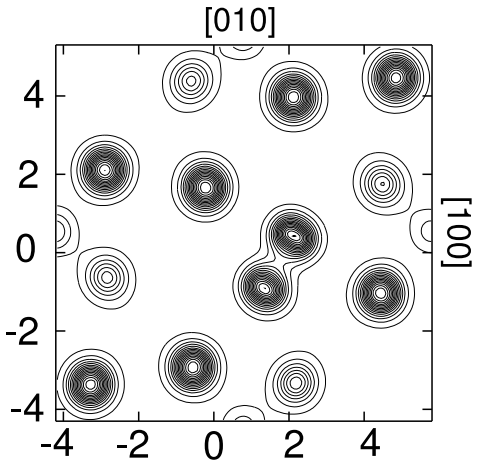


Figure 26: Charge density plot showing the arrangement of the oxygen dumbbell interstitial ($O_{i,db}$) within an oxygen (100) plane.

remarkable stability of this defect is caused by the binding σ_{p-p} states, which are located about one 1 eV below the VB.

Indium vacancies

In comparison with oxygen interstitials, indium vacancies play generally a minor role as an acceptor. Especially at lower Fermi energies the formation energy is extremely high (>5 eV). For highly n-doped conditions like in ITO, a transition could, however, occur from a donor compensation by oxygen interstitials to compensation by indium vacancies. This is possible since the favored charge states for indium vacancies ($q = -3$) and oxygen interstitials ($q = -2$) are different at high Fermi energies. The energetic ordering may further depend on the inaccuracies due to the XC-functional and the finite size corrections. The indium vacancy has transition levels to lower charge states throughout the band gap, but the formation energies attain high values for those Fermi levels so that their existence is less relevant. Energetically, the difference between vacancies on the two distinct indium sites (V_{In-b} and V_{In-b}) is less than 0.1 eV for all charge states. For both configurations a strong outward relaxation (13%) of the six neighboring oxygen can be found with a strong localization of the negative charge on the neighboring oxygen atoms.

5.4.3 *Defect associates*

Although In_2O_3 is highly ionic [223] indium anti-site configurations exhibit relatively low formation energies especially as compared to the oxygen anti-sites. It is a defect which can cause oxygen deficiency and n-type conductivity.

Stable anti-sites were also reported for SnO_2 and were explained by the multivalence of Sn [126]. Since indium possesses a lower oxidation number, which manifests for example in the more oxygen deficient compound In_2O, a similar behavior is plausible. Energetically, the indium anti-site (In_O) is very close to the interstitial and with comparable transition energies. The relaxation behavior of the defect is, however, different. In the case of the In_O, the relaxation process is highly non-symmetric. The energy minimum for this defect configuration is heavily displaced from the regular oxygen site resulting in an indium interstitial-oxygen vacancy defect pair. Surprisingly, the minimum energy is not at a regular interstitial c-position but on a b-site interstitial.

The binding energy (E_a) of V_O–In_i is relatively low (~ -0.3 eV) especially in contrast with anti-sites in SnO_2(-3.2 eV [126]). The binding energy is Fermi energy dependent and attains its minimum at the CBM.

Finally, the association of oxygen vacancies to di-vacancies $[V_O - V_O]$ was studied. Several configurations with distinct nearest neighbor configurations can be constructed, which practically do not differ in their formation energies. The Fermi level dependence of the formation energy and binding energy for one representative di-vacancy configuration is included in Fig. 27. The binding energy shows again a small, nevertheless negative value at the CBM.

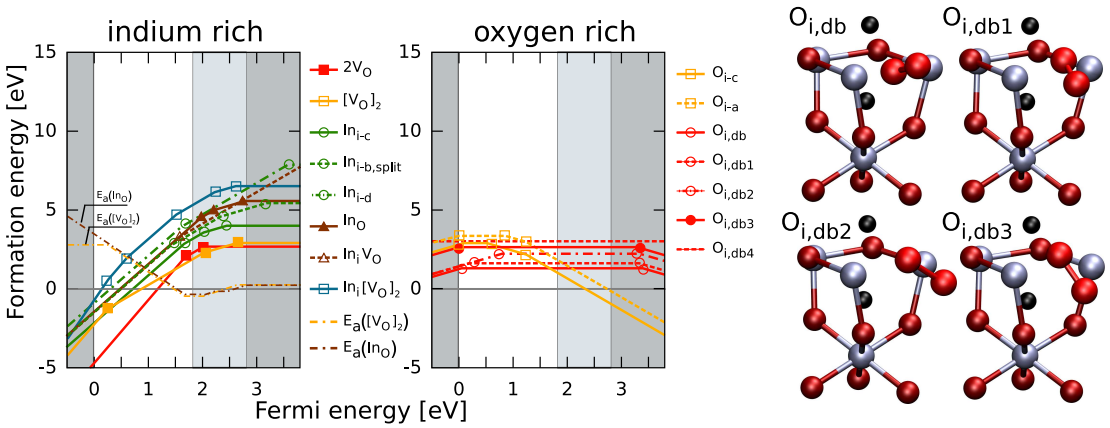


Figure 27: (left) Dependence of formation energy on Fermi level for defects and defect complexes. Here, the formation energy of two separated oxygen vacancies is compared with the formation energy of a divacancy. For Fermi energies higher than 1.5 eV there is a small negative association energy (E_a), whereas there is increasing repulsion at lower Fermi energy. A similar behavior is found for the anti-site (In_O). Indium *b*-site split-interstitials and *d*-site interstitials have slightly higher formation energies compared with *c*-site interstitials. (middle) Formation energies of oxygen interstitial defects are shown in the oxygen rich limit. Throughout the whole calculated band gap oxygen dumbbell defects are the major defect types. Four of these defects ($O_{i,db}$, $O_{i,db1}$, $O_{i,db2}$, $O_{i,db3}$) have lower formation energies than regular oxygen interstitials on *c*-/*a*-sites and differ mainly by the orientation of the dimer axis.

5.5 FORMATION ENTROPIES

5.5.1 *General aspects of entropy calculations*

Although atomistic calculations of vibrational contributions are possible, they are computationally demanding, especially if first-principles methods are employed. This is because the dynamical matrix for the ideal as well as the defective material has to be calculated at different cell sizes [224]. For treating elastic contributions to the defect formation entropy, different approaches like the embedded cluster method [228], the Greens function method [229] as well as the supercell method [224] were devised and tested. In this study, we have employed the plane wave pseudo-potential formalism for the force calculations in conjunction with the GGA [A2]. This method is expected to result in more reliable forces and is furthermore able to reproduce accurately the changes of force constants in the vicinity of the defect. Especially, the latter is difficult with other approaches.

In the past, materials under consideration included metals like Cu and Fe [224, 229], as well as Si [230, 231] but also ionic solids like AgCl or fluorite [228, 227]. Very little is known, however, about defect formation entropies in functional metal oxides.

Phonon density of states

Phonon calculations were carried out using the finite difference method described in Sect. 4.4. According to the symmetry of the system, finite displacive perturbations were created and the electronic ground state was calculated. The Hellman-Feynman forces were extracted for all atoms in order to build up the interatomic force constant matrix, which then allowed to construct the dynamical matrix at several $\boldsymbol{k}$-points. The diagonalization procedure was carried out using the fully coupled problem within the full harmonic approximation (FHA) [179, 236]. The phonon density of states was sampled using a Γ centered $15 \times 15 \times 15$ Monkhorst-Pack grid, which provided a convergence of better than $10^{-3} k_B$/atom in the calculated entropies. The displacements for the calculation of the force constant matrix were carefully checked to reside in the harmonic regime. The convergence of all parameters with respect to the frequencies of phonon branches was better than 0.1 THz [A2]. Finite size effects were assessed using two cell sizes (40 and 80 atoms) of necessarily two different geometries [68] (rombohedral and cubic). Finally, we note that the effect of LO/TO splitting due to long-ranged electrostatic forces at the Γ-point was not considered. The volume fraction around the Γ-point which may be affected is small compared to the entire Brillouin-zone volume and not important for thermodynamic calculations.

Vibrational entropy of the bulk

As a benchmark for our thermodynamic calculations we compare the calculated total entropy (Eq. 4.32) of the material with data from thermochemical tables [237] (see Fig. 28). The agreement with experiment is excellent and the error is less than $0.25 k_B$ per atom in a temperature region from zero to 1600 K. The deviations at higher temperatures originate from anharmonic contributions to the entropy, which are not accounted for in this fully harmonic calculation.

Defect formation entropy

We now turn to the discussion of the defect formation entropy (ΔS_f), which is in case of In_2O_3 essentially given by the difference of vibrational entropy of the defective system and the ideal system calculated for an identical number of atoms:

$$\Delta S_f = S_{def}^{vib} - S_{id}^{vib}. \tag{5.1}$$

The degrees of freedom in the supercell change when defects are introduced. In reality this is not the case, since the reservoir for exchanged atoms also contributes to the total entropy. Therefore, the reference entropy of the ideal cell has to be corrected in order

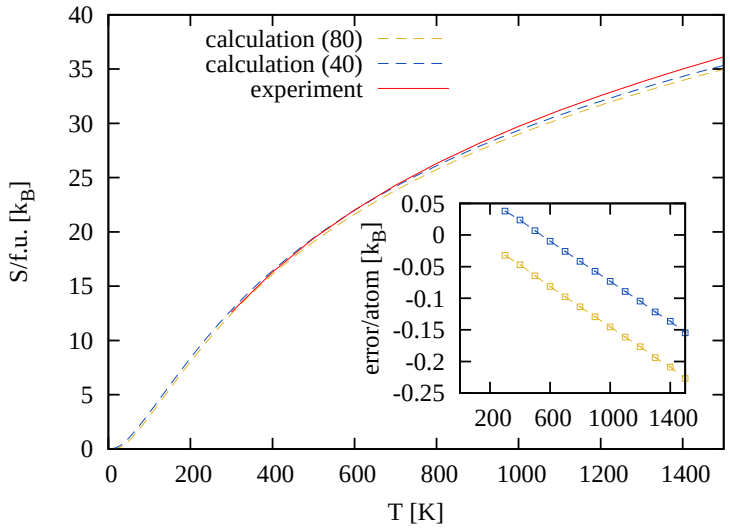


Figure 28: Comparison of vibrational entropy calculation with the experiment using two different cell sizes.

to match the number of atoms in the defective cell. In elemental materials, where all atoms contribute equally to the entropy of the entire system, this is simply achieved by rescaling the entropy of the ideal cell. Such a procedure, however, cannot be applied to a material like In_2O_3, where both the local bonding (octahedral coordination vs. tetrahedral coordination) as well as the different masses of the atoms ($\sim 16u$ for oxygen vs. $\sim 115u$ for indium) give rise to different entropy contributions from either constituents.

A more realistic estimate for the entropy contributions of the constituents, however, can be obtained by decomposing the phonon densities of states into partial phonon density of states (pDOS). Fig. 29 (left) shows the calculated partial phonon densities of states for indium and oxygen. It can clearly be seen that indium mainly contributes to the low frequency region and oxygen to the high frequency part. Fig. 29 (right) shows the corresponding partial entropy contributions as a function of temperature. Evidently, the vibrational entropy contributions of indium and oxygen are substantially different and, therefore, have to be distinguished if defect formation entropies are derived.

In the following, all data is taken relative to the partial entropies calculated from the partial density of states. This implies that only the lattice matrix would act as a source or sink for the defective species. If the material is in equilibrium with an external reservoirs like, *e.g.*, an oxygen atmosphere, the corresponding vibrational entropies of the O_2 molecule have to be taken into account.

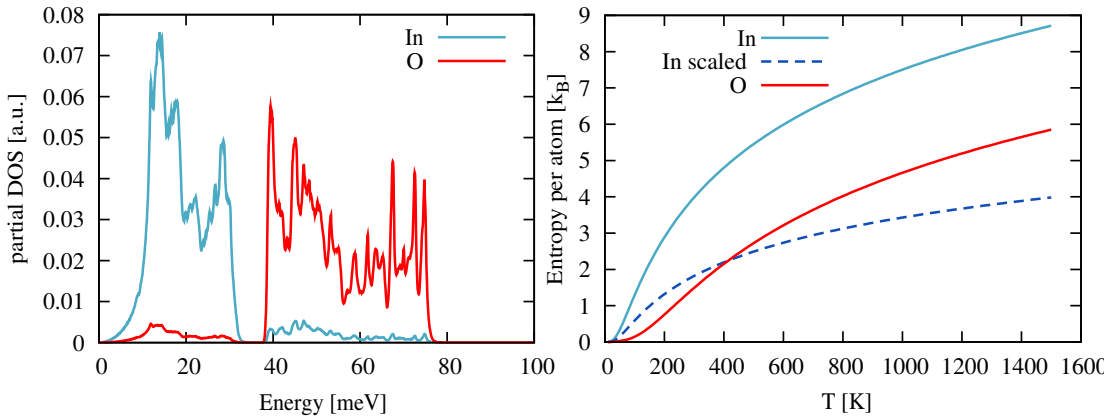


Figure 29: (left) The partial phonon density of states for indium and oxygen atoms. (right) Vibrational entropies of both constituents. Due to their higher mass the indium atoms tend to carry more entropy compared with oxygen. The dashed line represents a simpler approximation to the partial entropy. In this case the indium entropy is rescaled by the bond and mass ratios of the atoms.

5.5.2 *Elastic entropy contributions*

The formation of point defects leads not only to local atomic rearrangements but also to elastic distortions of the surrounding lattice. The defect formation entropy can, therefore, be divided into two major contributions, the core entropy ΔS_c and the elastic entropy ΔS_{el} [229, 230, 224]:

$$\Delta S_f = \Delta S_c + \Delta S_{el}. \tag{5.2}$$

The elastic displacement field around a point defect decays slowly ($\propto R^{-2}$) [206]. Thus, large system sizes are in principle needed to cover the elastic contributions. Mishin *et al.* [224], however, have shown in a detailed analysis of the elastic fields that finite size corrections to entropy are negligibly small if supercells are calculated at constant volume. Their values perfectly scale like $1/N$ (N is the number of atoms in the supercell) from a size of a hundred atoms on and even cells with 32 atoms still give reasonable results. Moreover, they show that extrapolating data from supercell calculations is by far more reliable than other methods, like the embedded-cluster approach [229]. Although only two different supercell sizes are accessible for In_2O_3, we will resort to this extrapolation scheme in the following section.

Supercell calculations are usually carried out at the volume of the ideal cell. In practice, however, one is actually more interested in the defect formation entropy at a

given (zero) pressure. The variation of entropy with volume can be simply derived from the following Maxwell relation

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial p}{\partial T}\right)_V = -V\left(\frac{\partial p}{\partial V}\right)_T \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_p = B_T\alpha, \tag{5.3}$$

with the isothermal bulk modulus B_T and the isobaric thermal expansion factor α. If we now consider the case of a system, which was initially held at constant volume with no displacements on the boundaries and is then relaxed to zero pressure conditions corresponding to stress free boundaries, the resulting entropy difference ΔS can be calculated from Eq. 5.3 as

$$\Delta S_f(V(p=0)) = \Delta S_f(V_o) + \underbrace{\alpha B_T \Delta V_{rel}}_{\Delta S}, \tag{5.4}$$

where the relaxation volume of the defective cell is $V_{rel} = V(p=0) - V_o$. Using this relation it is sufficient to carry out calculations at constant volume and to add the elastic corrections afterwards, given that the product of thermal expansion coefficient and isothermal bulk modulus as well as the relaxation volume are known.

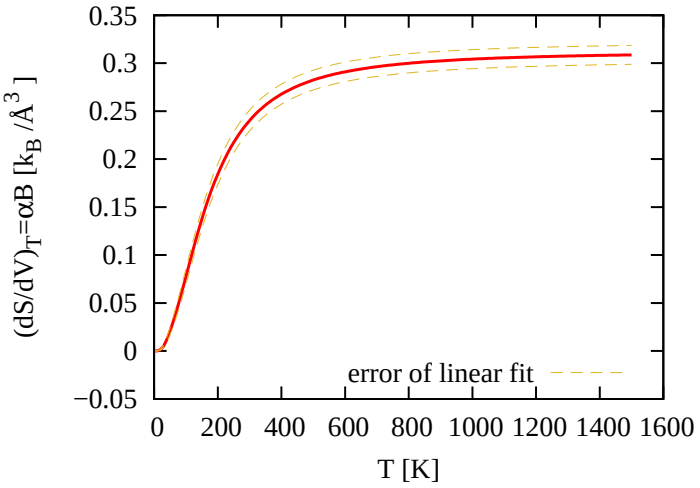


Figure 30: Product of thermal volume expansion coefficient and isothermal bulk modulus calculated from numerical derivatives of $S(V,T)$ with respect to volume at various temperatures.

The first value can be consistently derived from calculations within the quasi-harmonic approximation. By performing phonon calculations at several volumes and fitting polynomials of third order to the data at different temperatures, we have derived the temperature dependence of $(\partial S/\partial V)_T = \alpha B_T$. Fig. 30 shows the calculated results along with an error estimation obtained from the slope fitting error at each

temperature. The second quantity, namely the relaxation volume ΔV_{rel}, is readily available by relaxing the supercells to zero pressure conditions.

5.5.3 *Defect formation entropies at constant volume*

Firstly, we have calculated the defect entropies at constant volume by taking the entropy differences between defective and ideal cell of identical size. Entropy contributions of the missing/additional atoms were corrected by the values obtained from pDOS calculations as described before. We have considered the oxygen and indium vacancies in the charge states $q = +2$ and $q = -3$, as well as the oxygen and indium interstitial defects in charge states $q = -2$ and $q = +3$, respectively. The resulting data for T=1000 K are given in Table. 5 for two different cell sizes (40 and 80 atoms) in addition to the extrapolated value ($1/N \to 0$). Moreover, the calculated as well as extrapolated values are drawn in Fig. 31. The extrapolation leads to significantly larger values in the dilute limit for all defect types, which is conceivable because of the increasing elastic contributions. Mishin *et al.*[224] also found a characteristic upward trend using the

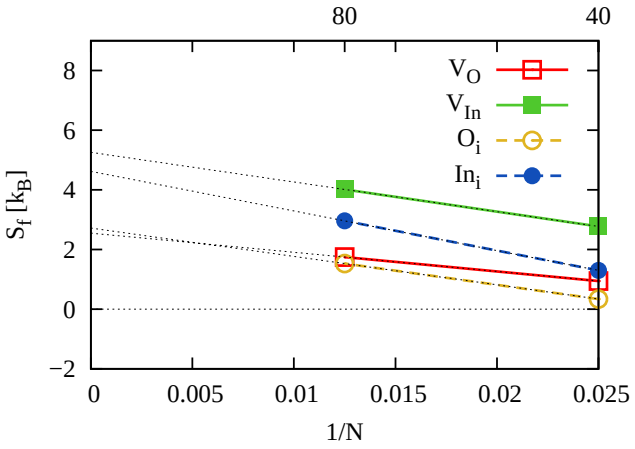


Figure 31: Calculated defect entropies for 40 and 80 atom supercells and their extrapolation ($1/N \to 0$).

Table 5: Defect formation entropies of four different intrinsic point defects in indium oxide at constant volume. The formation entropies are calculated at T=1000 K. The data contain results for 40 and 80 atom cells, as well as at the extrapolation limit.

	ΔS_f^{40} [k_B]	ΔS_f^{80} [k_B]	ΔS_f^{∞} [k_B]
V_O^{2+}	0.94	1.75	2.55
V_{In}^{3-}	2.77	4.02	5.26
O_i^{2-}	0.34	1.53	2.72
In_i^{3+}	1.30	2.96	4.62

extrapolation procedure, which was also independent of the defect character (sign of ΔV_{rel}). Since the extrapolation relies on only two data points, the error can be significant. A reasonable estimate for this uncertainty is the difference between the extrapolated entropies and the ones calculated for supercells with 40 atoms, which is about $2k_B$ for all defect types. As it will be shown below, this error has little impact on the calculated entropies at constant pressure, which are largely dominated by the dilatational part of the entropy ΔS_{el} in Eq.5.4 and can therefore be obtained rather accurately and consistently.

5.5.4 *Entropies at constant pressure*

Formation entropies at constant pressure ($p = 0$) are derived from the data given in Table 5 and the correction term ΔS described in section 5.5.2. Since the thermal expansion as well as the bulk modulus are positive numbers, the sign of the elastic correction for obtaining the formation entropies at constant pressure is only determined by the defect relaxation volume. We have calculated the relaxation volumes for each individual defect by iteratively scaling the volume until the residual pressure on the cell was less than 1 kBar.

By the sign of ΔV_{rel} the formation entropy may be increased or decreased. Defects which cause contraction of the lattice, like the oxygen vacancy and indium interstitial, decrease their formation entropy, whereas defects which expand the lattice produce more entropy. The effect is very pronounced and most significant for the acceptor-type defects. The results are summarized in Table 6.

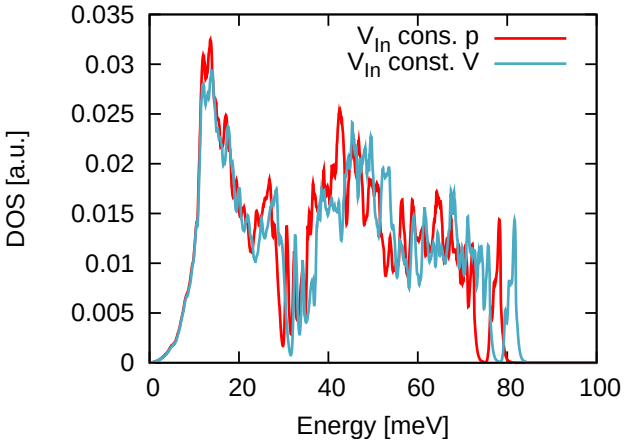


Figure 32: Phonon density of states for a defective cell containing an indium vacancy V_{In}^{3-}. Shown are the data of the constant volume and zero pressure calculation.

As a test case, we repeated the phonon calculation for the indium vacancy in a supercell relaxed to zero pressure and calculated the corresponding formation entropy. Fig. 32 shows the phonon density of states for both cases. It can clearly be seen that at zero pressure the vibrational spectra have homogeneously shifted to lower energy which is the reason for a larger formation entropy. Most interestingly, the formation entropy of 14.01 k_B obtained for the zero pressure case using a supercell of 80 atoms is

smaller than the number calculated by the extrapolation from constant volume data 16.39 k_B. This is in line with the predictions from elasticity theory, which suggest that the calculated entropy under zero pressure conditions becomes smaller with decreasing cell sizes [224]. Therefore, constant volume calculations together with elastic corrections are more appropriate for predicting the dilute limit at $p = 0$ than direct constant pressure calculations, at least if only small cell sizes can be accessed.

Table 6: Defect formation entropies for four intrinsic point defects of indium oxide at zero pressure and 1000K temperature. The first column contains the relaxation volumes of the defects in units of Å^3, whereas the second column describes the same quantity in units of the average atomic volume Ω. ΔS is the magnitude of the elastic correction term at 1000K. $\Delta S_f(V_0)$ is the calculated entropy at constant volume, $\Delta S_f(p = 0)$ the calculated entropy at zero pressure.

Defect	ΔV_{rel} [Å^3]	ΔV_{rel} [Ω]	$\Delta S_f^{\infty}(V_o)[\mathrm{k_B}]$	ΔS [k_B]	$\Delta S_f(p = 0)[k_B]$
V_O^{2+}	-17.5	-1.26	2.55	-5.37	-2.82
V_{In}^{3-}	36.5	2.63	5.26	11.13	16.39
O_i^{2-}	27.8	2.01	2.72	8.46	11.18
In_i^{3+}	-17.7	-1.27	4.62	-5.35	-0.73

Finally, in Fig. 33 we show in comparison the temperature dependencies of the formation entropies for constant volume and constant pressure conditions. As can be seen from Fig. 33 the high temperature approximation is well satisfied for the constant volume case, and the defect formation entropies for all four defects attain their final values at about room temperature. The high temperature limit, however, is approached at much higher temperatures in the constant pressure ($p = 0$) regime and varies by about $1k_B$ beyond room temperature. This effect has rarely been investigated previously [226]. Since the constant pressure case is more relevant for practical purposes, the high-temperature approximations should be handled with care.

5.5.5 *Discussion*

There is an important implication of our findings, which is specific to indium oxide. The formation energy is slightly larger for V_{In}^{3-} when compared with O_i^{2-} at about the doping limit of the material (section 5.3). The consequence is that the formation entropies can have a decisive effect on the defect's prevalence. In contrast to the energy, the formation entropy is higher for V_{In}^{3-}. This means that at high temperatures their relative stability could be reversed. As a matter of fact, two different acceptor defects have been identified at the doping limit of ITO. These defects were conjectured to be composed only of O_i^{2-} in association with the Sn dopants [109, 112, 105, 233]. Our

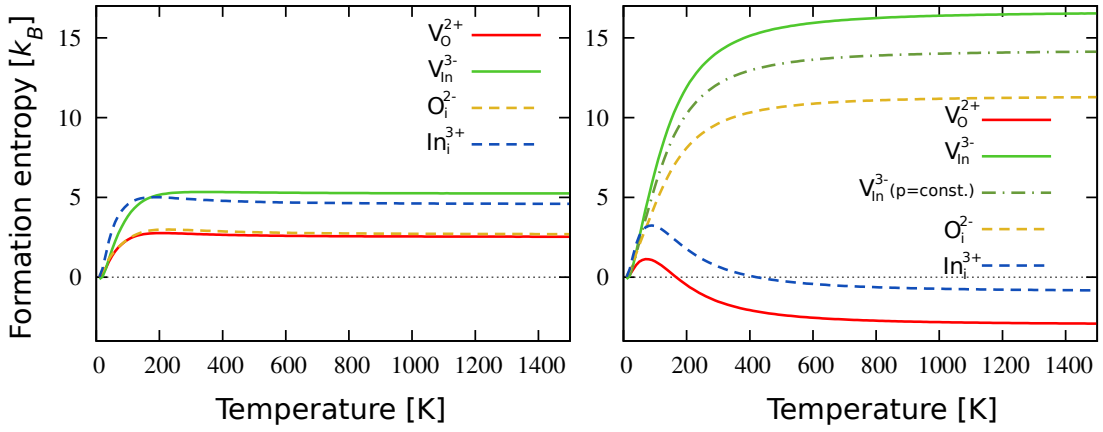


Figure 33: The temperature dependence of defect formation entropies. For the case of constant volume (left) the entropy values are stationary from room temperature on, whereas there is a significant temperature dependence beyond room temperature for the constant pressure case (right) due to the temperature dependence of $B\alpha$. For the case of the indium vacancy data obtained from calculations with pressure control of the supercell are also shown for comparison.

results give evidence that V_{In}^{3-} could be the second defect beside O_i^{2-} especially at high temperatures and partly stabilized by the defect entropy.

In contrast to acceptors, the formation entropies of the donor defects are negative and remain small (within the estimated error of our extrapolation procedure). As a result, entropy contributions do not change the donor defect equilibria, qualitatively.

5.6 SUMMARY AND CONCLUSION

By means of first-principles calculations within the plane wave pseudopotential formalism of density-functional theory we have conducted an extensive study of various possible intrinsic point defect configurations in In_2O_3. In the case of donor defects, oxygen vacancies as well as the indium interstitials are in principle capable of producing electrons within the conduction band, because positive charge states are favorable for Fermi energies over the whole calculated band gap. Our calculations reveal that the oxygen vacancy is more likely to produce *n*-type behavior than the indium interstitials, due to its lower formation energy. We identify a small binding energy between indium interstitials and oxygen vacancies. The absolute formation energies for indium interstitial related defects are, however, altogether high so that they can be excluded as potential intrinsic donors for In_2O_3. At low Fermi energies and oxygen rich conditions, oxygen interstitials in dumbbell geometries, equivalent to peroxo-ions, are identified as the dominant defect type in a large variety of orientations. These

oxygen dumbbells are characterized by their covalent oxygen bond. At higher Fermi energies, a transition from dumbbell-like configurations to doubly negative oxygen interstitials on symmetric *c*-sites occurs. O_i^{2-} is also the principal acceptor of In_2O_3 with, however, only small energetic difference to V_{In}^{3-} at high Fermi energies.

In addition to the formation energies, we have derived the formation entropies of point defects using a plane wave based electronic structure method in the full harmonic approximation and including elastic fields around the defect. As a general result our data provides evidence that entropy contributions are not always negligibly small and are in the case of the indium vacancy as large as $16k_B$. Corrections to the formation energy can, therefore, be 0.4 eV at room temperature and 1.5 eV at 1300K corresponding to a typical deposition temperature of thin films. Further, we found that defect formation entropies at constant pressure are largely determined by the relaxation volume. Acceptors and donors show a qualitatively different trend regarding their relaxation volume and have, therefore, different signs for ΔS_f. Furthermore, the magnitude is generally higher for acceptor defects due to strong outward relaxations, whereas the entropy of donors is small. It was also shown that the entropy contributions in the specific case of In_2O_3 are even able to change qualitative aspects of the defect equilibria.

6

THE ROLE OF XC-FUNCTIONAL: OXYGEN VACANCIES IN SNO_2, IN_2O_3 AND ZNO

6.1 INTRODUCTION

The results of the previous chapter indicated that oxygen vacancies have the lowest formation energies at reducing conditions. The results were obtained using the GGA+U functional which significantly underestimates the band gap. A prediction of absolute formation energies and electronic transition levels is only possible if results are corrected for this band gap error. Recalling section 2.2.5, there is strong evidence that oxygen vacancies are abundant defects and shallow donors in TCOs. The power-law dependence, $\sigma_{el} \sim p_{O_2}^{-1/6}$ of the electric conductivity on the oxygen partial pressure characteristic for V_O^{2+}, was observed in most TCOs [118, 114, 238]. Previous defect calculations could not satisfactorily explain this experimental finding as they were based on the LDA, GGA or LDA/GGA+U in the case of SnO_2 [132] and ZnO [239]. The previous chapter provided results based on the GGA+U for In_2O_3.

The results of different groups can be divided in two directions. Van de Walle and co-workers rely in their studies on the LDA/GGA+U techniques [240, 241] including an extrapolation method [239, 132, 134]. In this scheme the trends on formation energies of defects indicated by the application of the U parameter are extrapolated to the experimental band gap*. With this method they came to the conclusion that formation energies of oxygen vacancies are high in all TCOs and can, therefore, not contribute to the defect equilibria. Motivated by this finding, they proposed hydrogen incorporation as the major source of n-type conductivity in TCOs [242]. Alternative to this, Lany *et al.* use the GGA+U only for a correction of the VBM position and conduct defect calculations with GGA. They obtain low formation energies with, however, large ionization energies [125, 131]. On the basis of their analysis they propose an electron donation process which involves a persistent ionization of the oxygen vacancies by light described in (Sect. 2.2.5). Other explanations such as cation interstitials and donor complexes have also been proposed [243, 126]. Results from the previous chapter do, however, not support these views.

*This extrapolation was not used for In_2O_3 in the previous chapter.

The starting assumption of all these models is the validity of total energy differences in the LDA/GGA approximations. To resolve the controversy between present experiments and theory, it seems natural to probe the validity of the (semi-)local DFT methods used for calculating defect properties in TCOs. The failure of the LDA/GGA is primarily manifested as too small energy band gaps in comparison with experiments. Corrections are required in order to determine the positions of the defect charge transition levels with respect to the band edges and also the formation energies have to be modified, accordingly. The usual correction is to shift rigidly the band edges and/or the defect state itself (*e.g.* extrapolation methods) [239]. A close inspection of the magnitude of such corrections reveals that they are quite dramatic; the GGA band gap errors are $\sim$ 1.8 eV (64%), $\sim$ 3.0 eV (83%), and $\sim$ 2.7 eV (79%) for In_2O_3, SnO_2, and ZnO, respectively. A promising alternative approach is the replacement of LDA/GGA functionals by hybrid functionals [164, 165, 244, 166]. By including a part of the exchange energy in a nonlocal manner in self-consistent calculations they remedy the band gaps close to the experimental values (section 3.2.2).

Here, we test this methodology in a comparative study of the stability of oxygen vacancies in In_2O_3, SnO_2 and ZnO. In contrast to LDA/GGA($+U$) functional calculations the close agreement between the experimental and hybrid-functional band gaps does not call for further correction schemes resulting in a direct comparison of calculated vacancy formation energies and charge transition levels with the experimental ones. For In_2O_3 we have employed the HSE06 functional, whereas for SnO_2 and ZnO the PBE0 functional is used to reproduce the experimental band gaps. Note that these functionals are closely related and differ only by their screening length. Details on the computational setup and a comparison of calculated and experimental material data can be found in Refs. [A6] and [A8].

6.2 RESULTS AND DISCUSSION

6.2.1 *Formation energies*

The results shown in Fig. 34 reveal the existence of doubly positive (V_O^{2+}) and neutral (V_O^0) defects. Although the formation energies reported in literature vary strongly [196], it can clearly be seen in Fig.34 that the results obtained by using the LDA and GGA functionals are very similar. Both the formation energies as well as the charge transition levels ($0/+2$) are found at comparable values and an *n*-type behavior (the level ($0/+2$) is close to the bottom of the conduction band) is only found with respect to the calculated (too small) band gap. Also differences due to the inclusion/exclusion of the In/Sn *d*-states in the valence are always less than 0.1 eV in charge transition levels. Thus, the discrepancies in the reported values mainly result from different band gap correction schemes used.

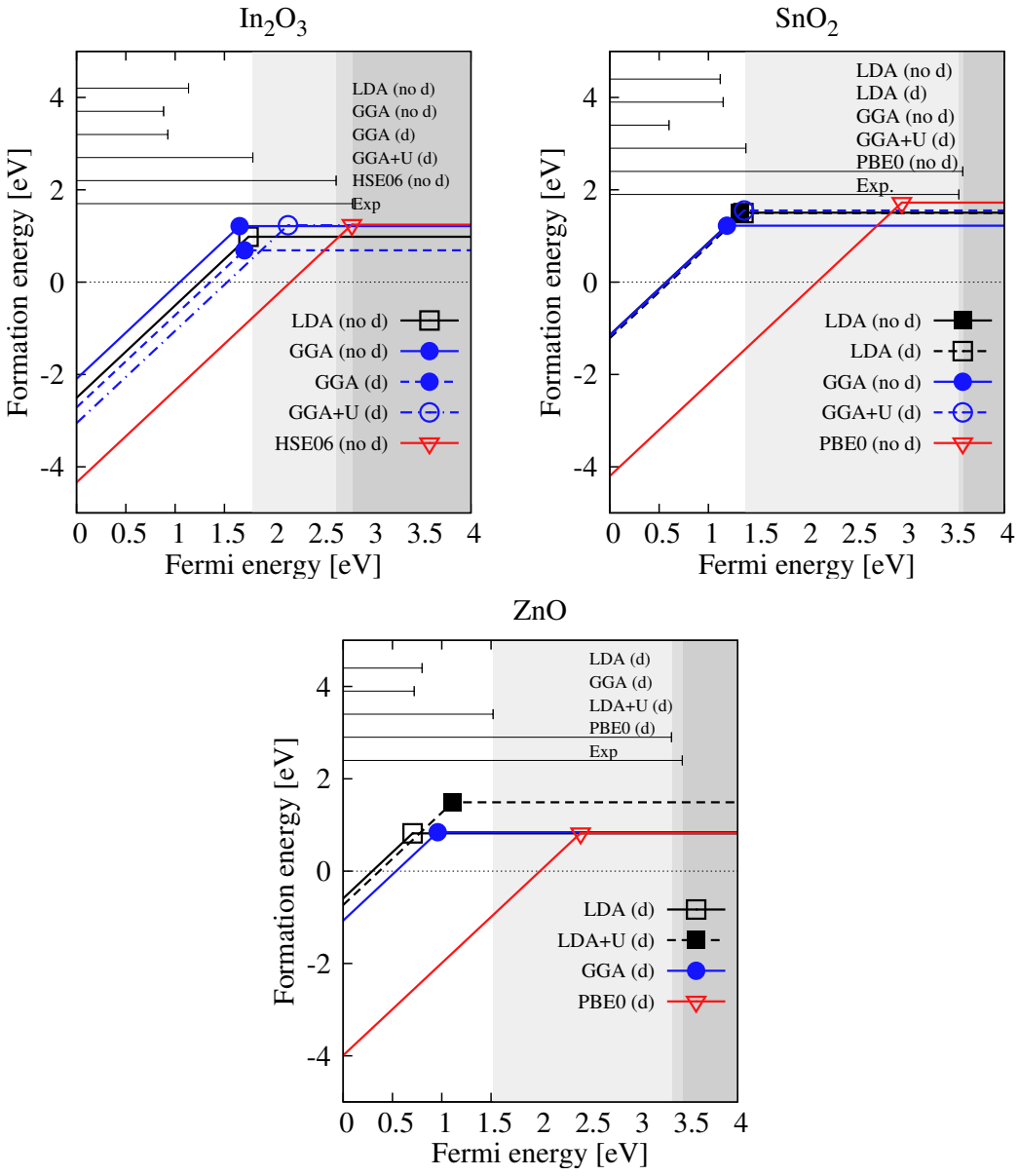


Figure 34: Formation energies of oxygen vacancies in In_2O_3, SnO_2 and ZnO (clockwise) at reducing conditions. The lines with positive and zero slopes refer to V_O^{2+} and V_O^0, respectively. The horizontal bars on the top show the calculated and experimental band gaps. The LDA/GGA+*U*, HSE06/PBE0 and experimental band gaps are denoted by shaded areas.

Applying the LDA/GGA+U method changes the results by giving slightly higher (0/+2) levels as well as higher formation energies for V_O^0. This is indicative for an upward shift of the defect state within the band gap. The shallow n-type behavior of In_2O_3 and SnO_2 is preserved but not that of ZnO. For V_O^0 the hybrid functionals give essentially the same formation energies as the LDA/GGA. In contrast, the transition levels (0/+2) are considerably pushed upwards, because the formation energies of V_O^{2+} decrease with respect to the LDA/GGA calculations. This is a remarkable result, since previous studies assumed more robust LDA/GGA formation energies for V_O^{2+} than for V_O^0, because the latter have no occupied defect states.

6.2.2 *Relaxation and band offsets*

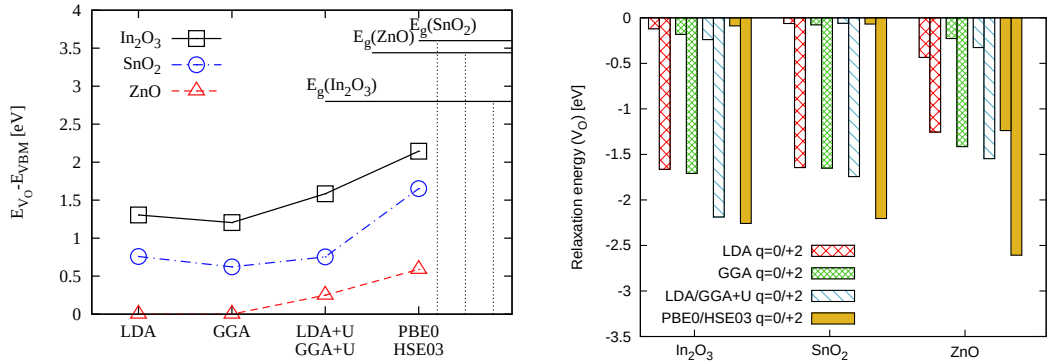


Figure 35: (left) Distance of the V_O^0 defect state from the VBM calculated with different functionals. (right) Relaxation energies for V_O in In_2O_3, SnO_2 and ZnO calculated with different functionals. In each case the left and right bars denote V_O^0 and V_O^{2+}, respectively.

In order to analyze the trends in the formation and charge transition levels we plot the relaxation energies for V_O^0 and V_O^{2+} as well as the single-particle defect state positions for V_O^0 (Fig. 35). The relaxation energy is defined with respect to the defect with atoms at their ideal lattice positions. Especially in In_2O_3 and SnO_2 the relaxation energies are remarkably larger for V_O^{2+} than for V_O^0. The relaxation energies increase in magnitude from the LDA/GGA values to those of LDA/GGA+U and then particularly strong to the hybrid functional values. The self-interaction correction inherent in the LDA/GGA+U and hybrid functional calculations increases the ionicity, and when a large ion relaxation occurs in the charged defect state, also the relaxation energy increases. This behavior of the relaxation energy explains only part of the differences between the exchange-correlation functionals in Fig. 34. A second effect is that the self-interaction correction lowers the VBM on the absolute energy scale [245], which favors the formation of positively charged defects with a further decrease of the

formation energy. Both reasons are, however, not effective for V_O^0 and their formation energies do not vary much (Fig.34). In contrast, LDA/GGA(+U) give clearly higher formation energies for V_O^{2+} than the hybrid functionals. For V_O^0 the differences are not so large. But, as shown in Fig. 35, the oxygen vacancy defect state rises for all materials when going from LDA/GGA to LDA/GGA+U. If extrapolation techniques or other band-gap corrections [239] are applied to the LDA/GGA+U results, the neutral-state formation energies would increase even more. This would lead to the conclusion that oxygen vacancies are not abundant in thermal equilibrium in these oxides.

Our results show that the band-offset for a scissor-like band-gap correction should mainly be attributed to the VB in contrast to previous studies. More importantly, we show that the relaxation energies predicted by the LDA/GGA are by more than 0.5 eV smaller than those calculated by hybrid functionals. This is an energy contribution which cannot be covered by band-edge shifting at all.

6.2.3 In_2O_3

On the basis of LDA/GGA calculations [131] the oxygen vacancy was conjectured to form a color center in In_2O_3. This has been obtained although the defect related oxygen vacancy state was located within the GGA CB. The assumptions which lead to this conclusion were that the energetic position of the vacancy state is independent of any band gap correction and that the band gap correction mainly acts on the CB states. As a consequence, the vacancy state was removed from the CB and a deep color center was predicted.

The hybrid functionals imply several differences. First, a major contribution to the band gap correction is VB related ($\sim$ 0.7 eV) and cannot increase the ionization energy of the defect as it was assumed by Lany and Zunger. The remaining correction is CB related and indicates the trend they predicted, however, less pronounced than anticipated. The defect state is removed from the CB but the thermal ionization level remains close to (in HSE06 even above) the calculated CBM. It is notable that the HSE06 band gap used for the calculations in In_2O_3 denote a lower bound for the measured gap [96]. After revision of the band gap model by Walsh *et al.* [99], several values have emerged ranging from 2.6 eV to 3.1 eV. The latter value is taken from HSE03+GW[†] calculations [97] whereas the lower value was obtained on epitaxial films [96] and from the evaluation of the "indirect"[‡] optical transitions in early experiments by Weiher [94].

With respect to the lowest band gap (2.6 eV) the oxygen vacancy should be readily ionized (Fig. 34), since the defect transition level is beyond the CBM in that case. The reason for this behavior is the large structural relaxation and energy gain of the defect upon removal of the electrons from the vacancy state. In this situation the

[†]Note, that the zero point vibration is not included in a usual GW calculation and the estimated relative accuracy of GW should not assumed to be better than 10% [246].

[‡]Prior to the revision of the band gap by Walsh *et al.* [99], optically forbidden transitions in In_2O_3 were attributed to an indirect band gap.

interpretations of Lany and Zunger [131] as well as those of Limpijumnong and van de Walle [134] cannot be sustained, because the ionization of the defect is immediate and the formation energies are relatively low as well. As it was pointed out more recently, this result can, however, not be sustained when the larger band gaps are taken as a reference [247]. Nevertheless, at higher temperatures, *e.g.*, $400^{\circ}C$, this should not be a problem due to the band gap narrowing $1 \times 10^{-3}\text{eV}/K$ [248] exhibited by In_2O_3. These conditions would bring the CBM again close enough to the electronic transition energy of the vacancy which would again result in a thermal ionization of the defect.

We can, therefore, conclude that the oxygen vacancy in In_2O_3 is an intrinsic defect which can provide electrons and can be incorporated at sufficiently high concentrations in order to explain oxygen deficiency. It is the intrinsic origin of the *n*-type behavior in this material.

6.2.4 *SnO_2*

In the case of SnO_2, our results are again in sharp contrast to previous calculations [126, 132], which predicted deep donors more than 1.8 eV below the CBM. According to our hybrid-functional calculations, the *n*-type behavior is favored also for SnO_2. In comparison with In_2O_3, however, the formation energy is higher and the electronic transition energy remains deeper within the band gap (0.5 eV). This trend is in accordance with experiments which indicate that maximum intrinsic electron concentrations are lower in SnO_2 than in In_2O_3 samples [20, 114, 120].

Compared with the experimentally observed intrinsic bulk donor levels ($E_1 = 0.03$ eV,$E_2 = 0.15$ [114]), the charge transition level of the oxygen vacancy is, however, too deep. In experiments it is also reported that two ionization levels exist for the oxygen vacancy which is in contrast with our findings. In addition, calculations of spin resonance parameters of V_O^+ suggest that the experimental assignments might not be valid [249]. Finally, we point out that the HSE type of calculations systematically underestimate the dielectric constants ϵ_r of the materials [246]. The PBE0 functional used for SnO_2 includes the Hartree-Fock exchange with a maximum screening length so that this effect is more pronounced. Since a low screening is crucial for the formation of color centers, the ionization energy of the vacancy could still be overestimated.

6.2.5 *ZnO*

For ZnO we find that the oxygen vacancy is a color center irrespective of uncertainties of the calculation. This is in agreement with previous studies [200]. In comparison with the LDA results [239] the transition level (0/+2) is nevertheless shifted towards the CBM by ~ 1.5 eV because the formation energy of V_O^{2+} decreases by more than 3 eV. The formation energy of V_O^0 is low, less than 1 eV, which naturally explains the non-stoichiometry. Only an insufficient fraction of the electrons will become, however, excited at relevant temperatures in order to explain the high electrical conductivity.

Therefore, it seems plausible to assume that also other mechanisms may contribute to the electrical conductivity like, *e.g.*, the persistent photoconductivity model (PPC) by Lany [125]. According to experiments, the photoconductivity decay time is long in ZnO [250], which makes a high average fraction of ionized vacancies [131, 125] possible under illumination. This explanation is supported also by the present study which predicts a large ionic relaxation for both, V_O^{2+} and V_O^0 with respect to the unrelaxed host. This is different for the In_2O_3 and SnO_2, where a substantial relaxation energy is only obtained for V_O^{2+}.

6.2.6 *Metastability and deep states*

An interesting feature of all *n*-type TCO materials is their deep states and structural relaxation behavior. The relaxation state of the vacancies is strongly interrelated with the defect state position. This observation forms the basis of the PPC model [125, 131]. In Fig. 36 it is illustrated for SnO_2, whereas it was reported in the literature for ZnO [125] and In_2O_3 [131] previously.

The figure shows the electronic DOS for the ideal material. In addition, the the V_O state in its two charge states is included as well. The single particle state of the neutral vacancy state is deep within the band gap whereas the contributions of the vacancy in the positive charge state appear close to the CB for V_O^{2+}. This behavior is distinct from that of typical color centers (*e.g.* MgO), because in the latter case the defect state position is independent of the relaxation state. The reason for this peculiar behavior is probably related to the intermediate position of TCOs between ionic and covalent materials. Color centers with deep states are characteristic for ionic materials with large band gaps. Therefore, traditional color centers are well described as confined electrons within a Madelung potential [206]. In contrast, *n*-type TCOs have a much smaller band gap and a stronger screening. Hence, the vacancy states are less localized but overlap with the nearest neighboring cations. The vacancy state is not confined in a strong potential, but is better described as a metallic bond between the nearest cation neighbors. With this in mind, the behavior of the vacancy state is plausible. When the vacancy relaxes outward upon electron removal from the vacancy state, the overlap between the cation orbitals is reduced and the energy of the state is lifted (Fig. 36). In the case of a neutral oxygen vacancy, the distance dependent vacancy state position

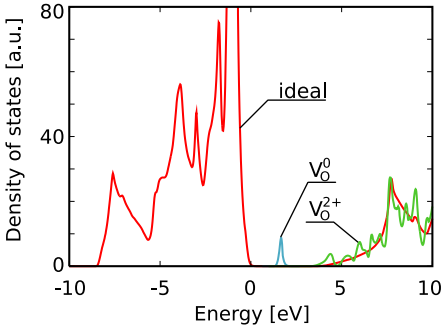


Figure 36: Oxygen vacancy states in SnO_2

can lead to a more persistent ionization of the defect, because the ionized vacancy state does not display a trapping behavior anymore (PPC) [125, 131].

While this interesting behavior is a general finding for all TCOs, we point out that in DFT calculations it can also be artificial and related to a finite size effect in some cases. In Fig. 37 we show a reaction coordinate describing the cation-cation interatomic distance of the indium nearest to the vacancy as a function of cell size for the two charge states +2 and 0. Within small supercells the relaxation behavior differs for the two charge states as described by Lany and Zunger. In contrast, the relaxation of V_O^0 diminishes as a function of cell size and vanishes for the largest supercell (640 atoms). This finding indicates that at the LDA/GGA level both charge states in fact resemble the same state of the system. A charge density analysis confirms the electron delocalization in the case of large supercells.

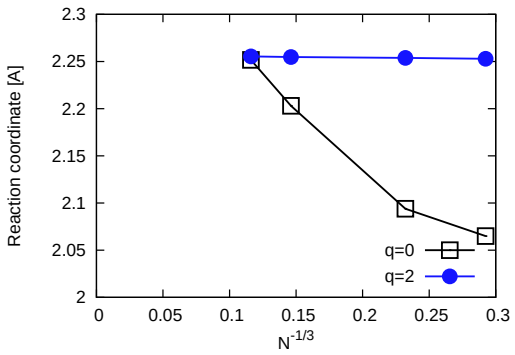


Figure 37: The relaxation strength of the oxygen vacancy in two different charge states as a function of cell size. At large cell sizes the relaxation of both charge states is equal.

To what extent this PPC effect can contribute to the conductivity of ZnO and SnO_2, should carefully be analyzed in the future. Due to the participation of excited states and a strong coupling of the phonon end electron system, this issue is beyond the scope of ground state DFT methods.

6.2.7 *Non-stoichiometry*

Some In_2O_3 samples, especially those which were vacuum annealed [251] or produced by thermal evaporation [252], can exhibit extreme electron concentrations and non-stoichiometry on the order of $1 \times 10^{20} cm^{-3}$ (In_2O_3). It is important to point out that the calculated formation energy for the oxygen vacancy cannot explain such concentrations under equilibrium conditions. Even smaller concentrations such as $1 \times 10^{17} cm^{-3}$ in the case of In_2O_3 are difficult to explain.

This is apparent when considering the phase diagram in Fig. 38. The isoline plotted for a chemical potential of $\mu(O_2) = 6.0$ eV denotes approximately the oxygen poor stability limit for all these oxides[§]. Even at annealing temperatures around 1000K, which are rarely used in experiment, the partial pressure of oxygen should be as low as 10^{-20}Pa in order to arrive at the oxygen-poor / metal-rich conditions. For this discrepancy there are several possible explanations:

[§]This is valid when the entropy contributions of solid phases are assumed to cancel.

(i) The calculation of the absolute formation energies are still in error. Additional errors arise from the heats of formation of the compound. This number enters in the calculation of the formation energies and is usually not reliable.

(ii) The oxygen vacancy is not the predominant defect causing conductivity. This could be possible when additional molecular hydrogen is present [114, 253]. Other intrinsic defects were excluded in the previous chapter for In_2O_3[¶].

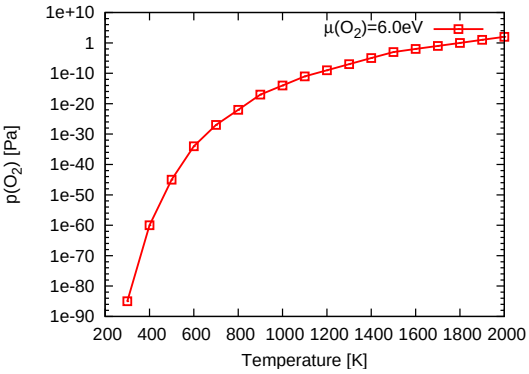


Figure 38: Oxygen-poor conditions translated into oxygen partial pressure and temperature.

(iii) The material is not in equilibrium. This is likely the case for thermally evaporated films, because in this case control is established via the concentrations not via the activity of the components. In the case of In_2O_3 for example, electrical measurements have shown a rapid increase of conductivity at high temperatures ($1600°C$). It is unlikely that this is due to an enhanced ionization of the vacancies as it was assumed by Lany and Zunger [131] previously. Because of the rapid weight loss measured by thermogravimetry and mass spectroscopy [120], this effect is associated with the decomposition and the material's stability limit. At the decomposition temperature, the material is brought into a non-equilibrium situation. Since the oxygen vacancy is the majority point defect it is plausible that initially the decomposition reaction is mediated by oxygen vacancy production.

(iv) A contribution to the conductivity arises from the surfaces of the material. Measurements are frequently conducted using thin samples with a large number of grain boundaries. As it will be shown in later chapters, electron accumulation layers are possible at surfaces of TCOs. A similar behavior could be exhibited by grain boundaries. For example, a thermogravimetric study of Mizusaki *et al.* [123] showed that a large portion of the released oxygen originated from the surfaces of SnO_2 at reducing conditions. They found that the bulk non-stoichiometry remained lower compared to the surface ($\sim 10^{-5}$) and in better agreement with the calculation.

(v) The actual oxygen pressure is lower when compared with the device reading. This is generally the case if additional reducing gases *e.g.* H_2/CO are present within the chamber. This explanation is most likely correct especially in the case of vacuum annealed samples.

[¶]The same behavior is obtained also for SnO_2 in the following chapter.

6.2.8 *p-type In_2O_3, SnO_2 and ZnO*

Our results have a general implication for all the studied TCO materials regarding the possibility of *p*-type doping. Namely, the formation energies predicted by hybrid-functional calculations are so low that also for oxygen-rich conditions the oxygen vacancy formation energies become negative for electron chemical potentials even above the VBM. This indicates that *p*-type doping is impossible for these materials close to thermodynamic equilibrium.

6.2.9 *Closer inspection of the LDA/GGA+U method*

Finally, we analyze the impact of the semi-core *d* states on the valence band structure of In_2O_3 and SnO_2. This is important because it is the physical motivation for the application of the LDA/GGA+*U* method. The LDA/GGA+*U* correction opens the band gaps by affecting on the *d*-state positions. Our hybrid-functional calculations produce band gaps in a reasonable agreement with experiments for both materials, *i.e.*, they are $\sim$3.6 eV and $\sim$2.6 eV for SnO_2 and In_2O_3, respectively.

Fig. 39 shows the calculated electronic densities of states for SnO_2, and In_2O_3 using GGA-PBE and hybrid functionals. The solid black lines always correspond to the situation where the *d*-states were not included into the valence whereas the dashed blue lines show the results for the *d*-states frozen into the core.

In comparison with GGA, the hybrid functionals shift the semi-core states downwards due to the reduced self-interaction errors similar to the GGA+*U* method [221, 132]. According to our hybrid functional calculations the band gap is not sensitive to the inclusion of *d*-states into the valence by more than 0.02 and 0.2 eV for SnO_2 and In_2O_3, respectively. We, therefore, conclude that the self-interaction of *d*-states cannot be a major reason for the band-gap error and suggest that the LDA/GGA+*U* method applied on *d*-states with extrapolations [132] is not a physically motivated approach to remedy the band gap problem.

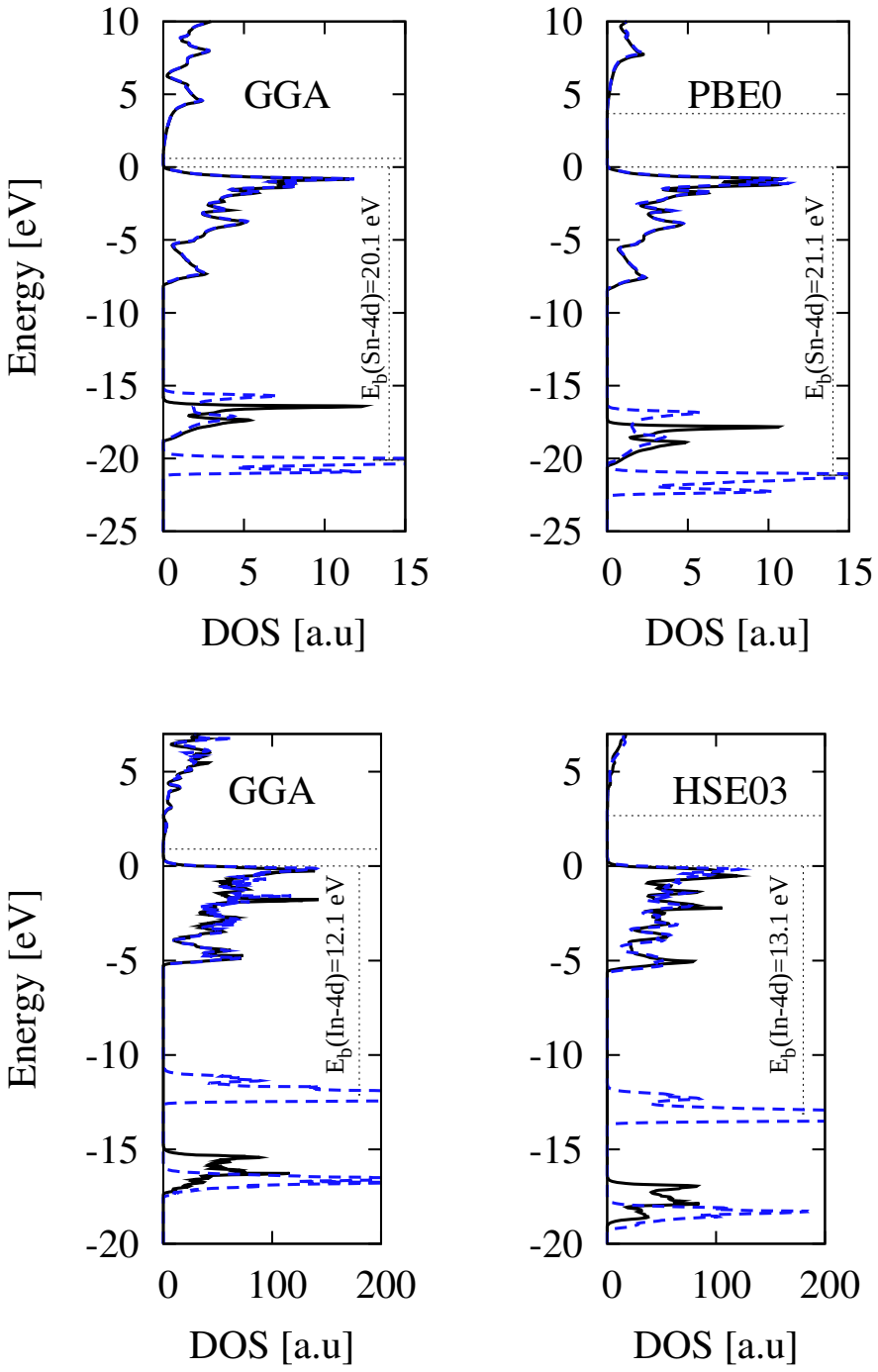


Figure 39: The electronic density of states for SnO_2(top) and In_2O_3(bottom). The GGA-PBE functional is compared with the hybrid functional results. Black solid and blue dashed lines correspond to calculations with and without *d*-states, respectively.

6.3 SUMMARY AND CONCLUSION

In summary, we conducted a comparative study on the intrinsic *n*-type conductivity of three TCO base materials In_2O_3, SnO_2 and ZnO. We applied the advanced hybrid-functional methodology beside the extensively used (semi-)local-DFT methodology. We identified major weaknesses of the standard LDA/GGA and tested conventional band-gap correction schemes in comparison with the hybrid PBE0/HSE06 functionals. We found that rigid shift corrections are not able to fully reproduce the hybrid-functional results. This is partly because the LDA/GGA underestimates severely, by more than 0.5 eV, the relaxation energies in vacancies. This underestimation cannot be covered by band edge shifting. According to our hybrid-functional calculations, the oxygen vacancy in In_2O_3 at reducing conditions is a shallow donor with a low formation energy. This strongly supports the view arising from experiments that abundant oxygen vacancies are the cause of the intrinsic *n*-type behavior. In comparison, the oxygen vacancies in SnO_2 show deeper transition energies and have higher formation energies. Generally, the propensity towards *n*-type behavior is significantly enhanced so that the overall agreement with experimental findings is improved by the hybrid functionals.

7

INTRINSIC POINT DEFECTS IN SNO_2: A HYBRID FUNCTIONAL STUDY

7.1 INTRODUCTION

After comparing different exchange correlation functionals, we further extend our investigations to other intrinsic defects in SnO_2. We, therefore, employ the PBE0 hybrid functional which we have tested in the previous chapter. This is, however, only possible with relatively small supercells (72 atoms). For the calculation of finite size effects we use LDA calculations. Details about the computational setup and the correction scheme can be found in Refs. [A8] and [A14].

Oxygen vacancies are of special interest and their properties were described in Sect. 6. In the following, other defects in SnO_2 are considered, analogous to the In_2O_3-study previously given in section 5. We include the cation vacancies, interstitials as well as donor defect pairs and the oxygen dumbbells. Oxygen anti-sites are not considered due to their extreme formation energies in In_2O_3.

7.2 FORMATION ENERGIES

In Fig. 40 we plot the calculated and finite size corrected formation energies for a series of defects (Sn_i, V_{Sn}, O_i, $O_{i,db}$, $[Sn_i - V_O]$ and $[V_O - V_O]$) as function of the Fermi energy and different oxygen chemical potentials. In addition, we have included the formation energies of V_O from the previous chapter 6. In Fig. 40 we also show the binding energies of defects consisting of more than one exchanged atomic species ($[Sn_i - V_O]$ and $[V_O - V_O]$) as a function of the Fermi energy (shown only for $\mu_O = 0.0$). Note, that binding energies are independent of the chemical potentials. Evidently, the binding energy has a dependence on Fermi energy due to the existence of varying charge states.

The defect equilibria of SnO_2 are largely dominated by donor-type defects. Within almost the entire range of oxygen chemical potentials, the oxygen vacancy has the lowest formation energy among all defects.

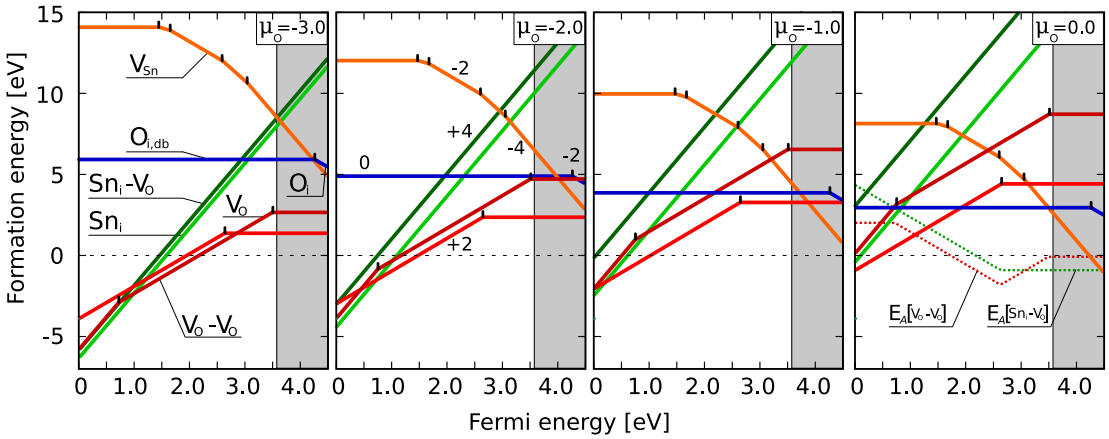


Figure 40: Defect formation energies of intrinsic point defects in SnO_2 as function of the Fermi energy. Shown are results calculated for four different oxygen chemical potentials indicated within the diagrams. Additionally, the binding energies E_A for two defect complexes are shown in the right panel.

7.2.1 *Donor defects*

Tin interstitials

Alternative donors to oxygen vacancies are the tin interstitials (Sn_i) which have relatively high formation energies at high Fermi energy. At lower Fermi energy they attain a negative formation energy similar to the oxygen vacancies.

For this defect our results massively differ from recent calculations by Singh *et al.* [132], where a semi-local exchange-correlation functional (GGA+U) was used in conjunction with extrapolations. We observe a formation energy for Sn_i^{4+} which is by $\sim$ 6 eV lower than the reported value, a trend which we have also observed for the oxygen vacancies (see chapter 6). In contrast to the non-local PBE0 XC-functional, the GGA+U method tends to increase the formation energies of donors. The resulting discrepancy is due to the exceptionally large band gap error in the case of SnO_2 and the high charge state of Sn_i magnifying the effect of the extrapolation procedure employed by Singh *et al.* [132].

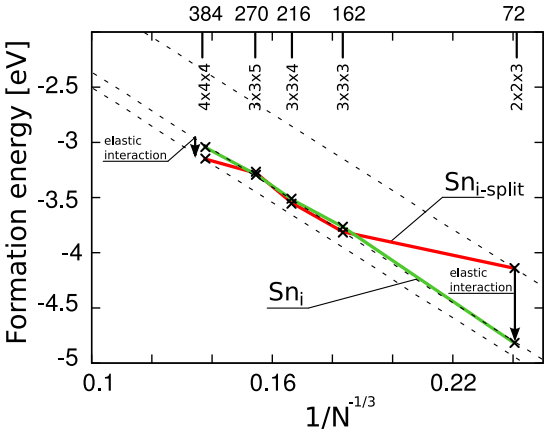


Figure 41: Scaling of the formation energy of Sn_i and $Sn_{i\text{-split}}$ defects as a function of the cell size.

Sn_i has only one charge state ($q = +4$) for all Fermi energies within the band gap. While it is well known that Sn^{2+} cations can exists, they cannot form in the highly symmetrical octahedral coordination of the interstitial site (Sn-5*s* lone-pairs mostly localize in asymmetric environments).

Due to the extreme charge state of the interstitial, strong electrostatic size effects are present. This is highly important for Sn_i because without finite size corrections (not shown) the interstitial is in fact more stable than the oxygen vacancy. This effect probably lead Kilic and Zunger to conclude that Sn_i is abundant in SnO_2 besides vacancies. If corrections are applied, however, the formation energies of the interstitials are shifted upwards by $E_{corr} = +4.20 \pm 0.07$ eV. The scaling behavior as a function of the cell size is shown in Fig. 41 and exhibits the expected behavior for the electrostatic image interaction $E \propto N^{-1/3}$.

While studying finite size effects, we have noticed another effect for Sn_i^{4+} which deserves attention. Using increasingly large supercells we found that the ground-state geometry of Sn_i^{4+} changes. While for smaller supercells (≤ 72 atoms) Sn_i^{4+} is located within the [001]-channels (Fig. 44), a split interstitial conformation ($Sn_{i\text{-split}}$) aligned in $\langle 100 \rangle$ direction is more favorable in larger supercells. This is consistent with the results of Godinho *et al.*[127] but also with results for interstitials in titania [254], a structurally equivalent material. In this geometry, the two participating cations are both located at the interstitial sites whereas one regular lattice site is vacated. Most appropriately, this defect is, therefore, described as a $[Sn_i - V_{Sn} - Sn_i]^{4+}$ defect complex due to the large spacial separation between the two interstitials. The reason for this transition with cell size is related to strain within the (001) plane caused by the interstitial. When comparing the scaling behavior with that of the regular interstitial, irregularities can be observed (marked by arrows in Fig. 41) which occur whenever the cell-dimension is increased perpendicular to the *c*-direction while the scaling behavior is identical when the *c*-direction is increased. Energetically, the two defects differ by less than 0.2 eV. The propensity towards the formation of split interstitials in rutile structures can be related to the sign of the electrostatic potential at the interstices (Fig. 44) which favor positive species.

Similar to other TCOs (*e.g.*, ZnO and In_2O_3) the cation interstitials are a minority shallow donor and have higher formation energies when compared with oxygen vacancies. This is in contradiction to the conclusions made in Ref. [126] where both Sn_i and V_O have equally low formation energies. It is also noteworthy that in contrast to titania, a structurally related material, there is no crossover in formation energy [254] of the oxygen vacancy and the interstitials. The higher stability of cation interstitials in titania can be explained by the more covalent nature of the material.

Interstitial-vacancy complexes

Because of the existence of two low energy donor defects, Kilic and Zunger [126] proposed that these defects could combine in order to form bound defect complexes $[Sn_i - V_O]$. These ideas were also investigated for other TCOs like ZnO [243] and for

In_2O_3 in chapter 5. In the case of ZnO this defect has been proposed as an additional source for *n*-type conductivity while for In_2O_3 we have concluded that the formation energies are too high to be responsible for non-stoichiometry or *n*-type conductivity.

More recent GGA+*U* calculations, however, did not support this idea [132] for SnO_2. Using the PBE0 functional, we find that the formation energy of the $[Sn_i - V_O]$ complex is again much lower than results obtained by semi-local exchange-correlation functionals and the binding energy of these complexes becomes negative at Fermi energies close to the CBM (Fig. 40). In fact, the binding of interstitials and vacancies is highest in SnO_2, when compared with In_2O_3 (Sect. 5.4.3) and ZnO [243]. Our calculated binding energy (~ -1 eV) is, however, considerably smaller when compared with previous reports (~ -3 eV [126]). The absolute values of formation energies remain nevertheless high, higher when compared to Sn_i^{4+} alone.

The defect complex $[Sn_i - V_O]$ occurs in several possible configurations. The lowest energy configuration we have identified differs from the one proposed by Kilic and Zunger [126]. This is because the lowest energy configuration is a vacancy connected with a $Sn_{i\text{-split}}$ instead of Sn_i in the regular interstitial position. Like Sn_i, these defect complexes assume the charge state $q = +4$. This arises from the ability of Sn to localize a lone pair when brought into an asymmetric and low coordinated environment. In the present case the transformation is probably also driven by the large amount of accumulated positive charge. The defect complex can be described as a $Sn_i^{2+} - V_O^{2+}$ defect associate with Sn_i localizing its lone-pair into the vacancy.

Double oxygen vacancy

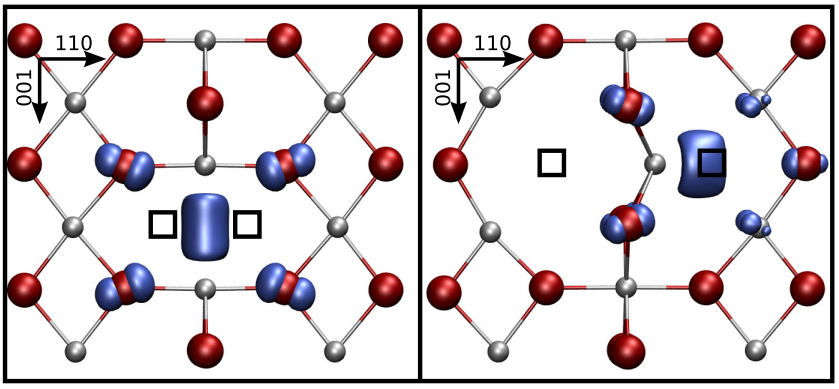


Figure 42: The two most stable double oxygen vacancy configurations. The charge density of the stabilizing states are drawn within the ball and stick models and the location of the two vacancies is indicated by boxes.

We have further explored the possibility of associations among donors. Since oxygen vacancies are the most abundant defects it is plausible to also study their interaction in several different configurations. As it is shown in Fig. 40 $[V_O - V_O]$ have the overall lowest formation energies under tin-rich conditions $\mu_O \sim -3.0$ eV. Fig.42 shows the two most stable arrangements of two oxygen vacancies both having $q = +2$ as the predominant charge.

The first configuration is an M-center [206] with two four-fold coordinated Sn atoms facing each other in close proximity (left side of Fig.42). The electron density is increased between the two atoms indicating an Sn-Sn bond. The corresponding bonding state is located close to the VBM in the electronic density of states whereas the anti-bonding state is located close to the CB explaining the predominant charge state $q = +2$. The resulting defect levels can simply be understood in terms of a di-hydrogen like bonding. It is notable that the $[V_O - V_O]$ is a shallower donor compared compared to single vacancies.

The second most stable configuration ($\Delta E = +0.51$ eV) is obtained by positioning oxygen vacancies on two opposite sides of a cation with a subsequent relaxation of the cation to either side (Fig. 42 (right)). As a result, a shallow vacancy is obtained on the one side whereas the cation localizes its lone-pair into the vacancy on the opposite side forming a deep state. The maximum association energy of the double vacancy can be as large as $E_A^{V_O - V_O} = -1.5\, eV$ with its maximum at a Fermi level position close to the CBM (see Fig. 40) where it is also usually observed in experiments [85]. For highly reduced samples this defect is important and an additional source of charge carriers and non-stoichiometry.

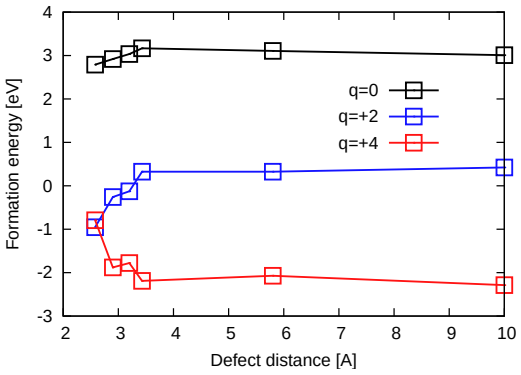


Figure 43: The formation energy of oxygen double vacancies as function of their separation. At short distances there is a series of different bound complexes.

Besides the defects shown in Fig. 42, there are several other bound $[V_O - V_O]$ configurations. The distance dependence of formation energies of $[V_O - V_O]$ is shown in Fig 43. In close proximity there is a linear energy gain in the charge states $q = +2$ and $q = 0$. In the case of $(V_O - V_O)^{4+}$ the interaction energy is positive due to the coulomb repulsion.

7.2.2 *Acceptor defects*

Tin vacancies

Among the different acceptors in SnO_2 (Fig. 40), tin vacancies are the most stable ones. In the past, the existence of Schottky-type equilibria has been discussed [255]. Based on our calculations, however, we can exclude the formation of V_{Sn} for intrinsic, acceptor, and even for moderately donor doped materials. Note, that even in the oxygen rich limiting case $\mu_{O_2} = 0$ eV and for Fermi energies at the CBM, the formation energy of V_{Sn}^{4-} is very high ($\sim$ 2.5 eV). Consequently, the Schottky defect reaction

$$0 \leftrightharpoons V_{Sn}^{4-} + 2V_O^{2+} \tag{7.1}$$

cannot be invoked to explain anomalies in the defect chemistry of SnO_2[255]. The calculated formation energy for this reaction is as large as $E_{7.1} = 14.5$ eV for the separated vacancies and can be safely ignored. In order to further strengthen this argument we have also considered bound forms of Schottky defects according to reactions

$$0 \leftrightharpoons [V_{Sn} - V_O]^{2-} + V_O^{2+} \tag{7.2}$$

$$0 \leftrightharpoons [V_{Sn} - 2V_O]^{0} \tag{7.3}$$

in several different geometries. The smallest formation energies are still high $E_{7.1} =$ 8.7 eV and $E_{7.2} = 8.1$ eV. We do not show the Fermi energy dependence since the associated defects assume the charge states given in equations (7.2) and (7.3) for Fermi energies throughout the entire band gap. In the latter case the formation energy is calculated by considering also the formation energy of one isolated oxygen vacancy. We further find that a $[V_{Sn} - V_O]^{2-}$ defect (reaction 7.2) would produce a Fermi pinning level (zero formation energy) at considerably higher Fermi energy when compared with the isolated tin vacancies. From the calculated values it is evident that the equilibrium concentration of Schottky defects will be low ($c_{Shottky} < 10^2 cm^{-3}$).

Oxygen interstitials

The formation energies of negatively charged oxygen interstitials are high due to the size of the oxygen anions on the one hand, but also due to the sign of the electrostatic potential at the interstices (Fig. 44) in SnO_2 on the other hand. The latter was calculated using simple Coulomb potential with idealized charges [256, 257]. Additionally, the *n*-type TCO materials generally exhibit high electron affinities (Sect. 2.2). This results in a deep position of the Fermi energy with respect to the vacuum level (high work function). In such cases, oxidation reactions are generally favored and can lead to peroxide formation in the lattice. Consequently, negative oxygen interstitials are disfavored with respect to dumbbell peroxide interstitial formation according to reaction

$$O_i^{2-} + {O_O}^0 \leftrightharpoons (O_2)_O^0 + 2e^-. \tag{7.4}$$

For several TCO materials $O_{i,db}$ defects were reported [235, 258, 131] including SnO_2 [127, 259]. In SnO_2, however, the special case is encountered where oxygen interstitials are almost exclusively present in this single form (Fig. 40). The transformation into O_i^{2-} is found only at a Fermi level within the CB. At this Fermi level V_{Sn}^{4-} attains a lower formation energy so that O_i^{2-} can never be a majority defect. The $O_{i,db}$ on the other hand is the predominant defect under highly oxidizing condition. However, even in this limiting case the formation energy is still > 2 eV.

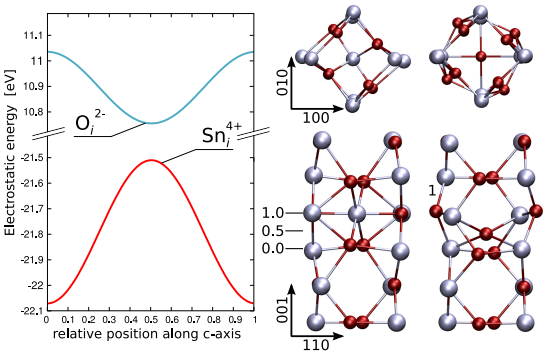


Figure 44: The electrostatic potential within the interstitial channels in Rutile along $\langle 001 \rangle$. (right) Arrangement of cations and anions in the interstitial channels.

7.3 SUMMARY AND CONCLUSION

We have revised the defect equilibria of SnO_2 using hybrid functional methodology. It was shown that the oxygen vacancies are the most abundant defects in this material. Further, cation interstitials have high formation energies, although the values are lower by 6 eV compared to previous studies. In this context we have presented the importance of finite cell size corrections for the Sn_i^{4+} with its extreme charge state.

Association energies among Sn_i and V_O were shown to be high but these defect pairs as such could be excluded for equilibrium conditions on the basis of high formation energies. On the other side, associations among V_O are important and lead to low energy donor defect pairs especially under reducing conditions.

The principal acceptor in SnO_2 is V_{Sn}^{4-} with high formation energies even in the oxygen rich limiting case. Negatively charged oxygen interstitials have a small stability range because they easily transform into $O_{i,db}$. At any external condition, oxygen interstitials are incorporated in this neutral form. Acceptor defects as well as neutral oxygen dumbbell defects are relevant only under extreme *n*-type and/or oxidizing conditions, respectively.

8

DEFECT EQUILIBRIA AND DOPING LIMITS: In_2O_3 VS. SnO_2

8.1 INTRODUCTION

We are now in the position to compare the defect energetics of the two materials SnO_2 and In_2O_3. The comparison is slightly complicated by the fact that for In_2O_3 the data was obtained using the GGA+U method whereas the data for SnO_2 is based on PBE0. Considering trends of the corrections which the hybrid functionals provided in the previous chapter, the data from chapter 5 can be "translated" to make comparison possible. According to the hybrid functionals, corrections to the the data of Fig. 24 should mainly comprise a downshift of the VB. The formation energies of donors decrease whereas those of acceptors increase. The formation energies of neutral defects remain approximately constant. These changes do not alter the energetic order of the defects in In_2O_3 so that a qualitative comparison is possible.

The major similarities are:

- The dominating defect at reducing conditions is the oxygen vacancy.
- Donor formation is generally favored over acceptor defect formation.
- Cation interstitials have always higher formation energies when compared with oxygen vacancies.
- Oxygen dumbbells are the majority defect under oxidizing conditions for the undoped host.
- Associations among donors are favorable.

More interesting are, however, the differences between the two materials:

- The V_O in SnO_2 is a deeper donor compared to vacancies in In_2O_3.
- The principal acceptor is O_i^{2-} for In_2O_3 whereas it is V_{Sn}^{4-} for SnO_2.
- Negative oxygen interstitials form only in In_2O_3. Oxygen interstitials in SnO_2 are present exclusively in neutral form ($O_{i,db}$).

- Association energies for donors are on the order of 0.3 eV in In_2O_3, whereas they are on the order of 1 eV − 1.5 eV for SnO_2.
- Double oxygen vacancies only form in SnO_2 under equilibrium conditions.
- The formation energies of all defect types are larger in SnO_2 when compared with In_2O_3, *i.e.*, SnO_2 is more resistant towards defect formation.

Especially the final point deserves special attention and has technological implications. This conclusion is not only valid when SnO_2 is compared with In_2O_3 but also with other TCOs. Table 7 shows the formation energies of several point defects in

Table 7: Defect formation energies of several defects in several TCOs. For $O_{i,db}$ and V_{cation} values correspond to oxygen-rich conditions whereas V_O formation energies are in the cation-rich regime. All numbers were obtained for charge neutral defects and with LDA.

Material	Formation energy		
	$O_{i,db}$	V_O	V_{cation}
ZnO	1.70	0.82	1.61
CdO	0.96*	0.34 *	1.03*
In_2O_3	1.03	0.98	5.14
Ga_2O_3	1.42	1.22	5.00
SnO_2	2.39†	1.50†	8.16†

*lowest value
†highest value

different TCOs. For the sake of comparison the numbers were obtained with LDA for all compounds from supercells of similar size ($\sim$ 100 atoms). SnO_2 exhibits the highest formation energies for all defects among all TCOs, whereas CdO shows the lowest. Moreover, this trend correlates neither with the band gap nor with the heat of formation.

A technologically important case where the relative formation energies are relevant is for doping limits. TCOs can be doped to a rather different level and the origin of the doping limits is not known in most cases. In this context In_2O_3 and SnO_2-based TCOs are a good example: ITO shows the highest conductivities of all the TCO materials [260, 261]. Some disadvantages are, however, connected with the uncertainties surrounding the indium supply (Sect. 2.2.7). ATO, on the other hand, is cheap and abundant but lower conductivities [138] hinder its commercial application. Doping SnO_2 on the anion site with fluorine (FTO) increases the conductivities, but FTO is still significantly outperformed by ITO [260]. Additionally, it is difficult to use

fluorine in conjunction with sputtering techniques whereas chemical vapor deposition processes lead to a undesirably high roughness.

In this concluding study, we revisit the problem of the thermodynamic stability of acceptor-type point defects in In_2O_3 and SnO_2 at high doping levels using hybrid-functional DFT in conjunction with finite-size corrections based on the local density approximation (LDA). The formation energies for SnO_2 were taken from chapter 7 whereas the formation energies for In_2O_3 were recalculated using the HSE06 hybrid functional. The aim is to quantify and compare the maximum free electron concentrations achievable in In_2O_3 and SnO_2 as a function of doping and oxygen partial pressure. For details about the methodological setup, the reader is referred to Ref. [A8] as well as chapter 6 and section 3.2 in the methodology part.

8.2 DOPING LIMITS

8.2.1 *Formation energies*

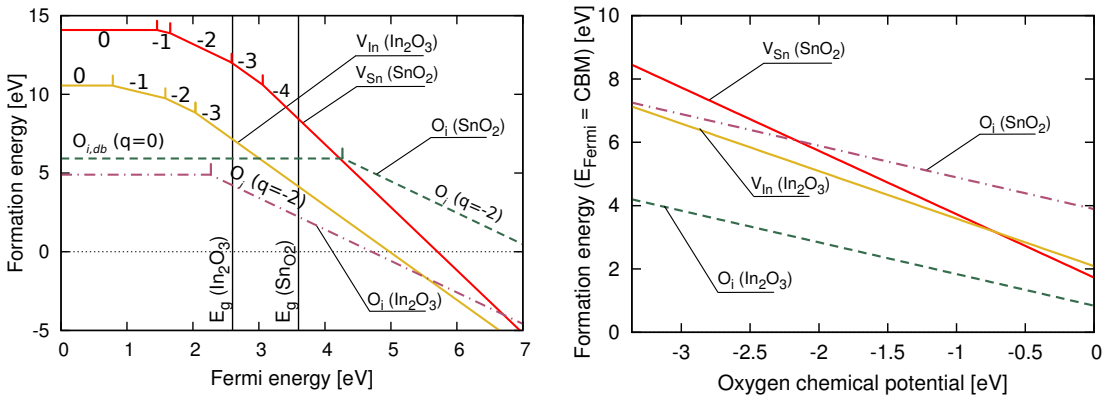


Figure 45: (left) Formation energies of acceptor defects in SnO_2, and In_2O_3 in the metal rich limit. For the two materials the doping limits are approximately given by the lowest-energy intersections of the formation energies with the zero energy line. In SnO_2 intrinsic compensation occurs at significantly higher Fermi energies.
(right) Formation energies of acceptor defects at the conduction band minima as a function of the oxygen chemical potential. The stability limits for the two materials are similar.

Fig. 45 shows the finite-size corrected formation energies of acceptor defects in In_2O_3 and SnO_2 for maximally reducing conditions. At this limit the formation energies of all acceptors attain their highest values and are suitable for the discussion of maximum doping limits. The doping limits are approximately given by the lowest energy intersections of the formation energies with the zero energy line in the left part of the figure. For both materials the intersections are significantly (> 2 eV) beyond

the conduction band minimum. Moreover, it is very clear that SnO_2 is intrinsically compensated only at Fermi energies considerably higher than In_2O_3. Namely, in this limit the compensation occurs above ~ 4.7 eV and ~ 5.7 eV for In_2O_3 and SnO_2, respectively.

In the right part of Fig.45 the formation energies of the acceptor defects are shown in their highest negative charge states for the Fermi level at the CBM as a function of the oxygen chemical potential. Thus, the *n*-dopability of both materials can be compared irrespective of the different band gaps. The figure illustrates the fact that unless degenerate doping is achieved, in both materials no acceptor defect can contribute significantly to the defect equilibria. For example, the formation energy of the double negative oxygen interstitial in In_2O_3 is ~ 1 eV in the oxidizing limit and for the Fermi energy at the CBM. Since the oxidizing limit is unrealistic at elevated temperatures and the Fermi energy is generally below the CBM for undoped and oxidized In_2O_3 samples [262, 263, 264], we can exclude the occurrence of intrinsic acceptor defects under any experimentally accessible conditions. As can be seen in the right part of Fig. 45 this effect is even more pronounced for SnO_2. For the Fermi energy at the CBM and at most oxidizing conditions the formation energies of both acceptors are larger than 1.8 eV. Note also that the use of the calculated heats of formation for the determination of the stability range underestimates the formation energies in the oxygen-rich limit whereas the values are more reliable in the metal-rich limit.[212]

8.2.2 *Electron concentration*

In order to illustrate this finding more clearly, we have calculated the electron concentrations self-consistently on the basis of the finite-size corrected formation energies of acceptor defects obtained from hybrid-DFT total energy calculations. The defect concentrations were obtained by the usual Boltzmann type expression,

$$c = c_0 \cdot \exp\left(\frac{-\Delta\Omega_f}{k_B T}\right), \tag{8.1}$$

where c_0 is the concentration of available sites for the defect and $\Delta\Omega_f$ the defect formation energy. Since we are presently only interested in the defect properties of the intrinsic acceptors, the *n*-dopant is assumed to be ideal in this calculation, i.e., it has the ionization probability of unity and it is ideally soluble. For both materials we use temperature-independent parabolic band edges with effective electron and hole masses of 0.3 m_e and 0.6 m_h, respectively [20]. Because the Fermi level can enter into the conduction band, we use the Fermi function instead of the Boltzmann approximation in order to integrate the charge carrier densities.

Fig. 46 shows the free electron concentration due to heavy *n*-type doping as a function of the oxygen partial pressure in the range from 10^{-15} to 10^5 Pa, which is accessible within experiments. The oxygen partial pressure is obtained from the oxygen chemical potential via the ideal gas law and using electrochemical tables

for the temperature of 600° C [212, 180]. The conductivities are given for three different doping concentrations of 2×10^{20} cm^{-3}, 6×10^{20} cm^{-3} and 2×10^{21} cm^{-3}, corresponding to the range of $\sim 1\% - 10\%$ of substituted cations. The free-electron

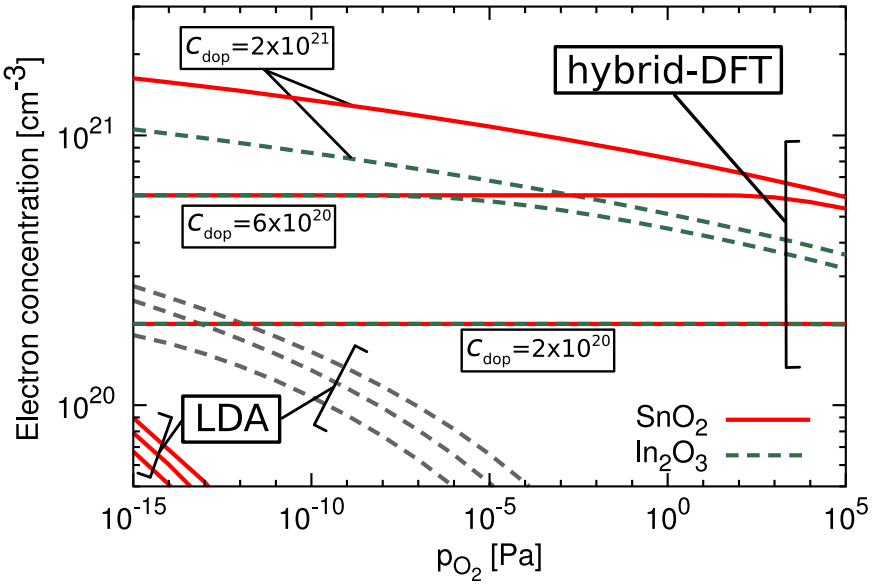


Figure 46: Free electron concentration in In_2O_3 and SnO_2 as a function of the oxygen partial pressure. The results are shown for a temperature of 600°C and for three different doping levels. For comparison, also the corresponding LDA results are given.

concentrations are highest at low oxygen pressures and mainly determined by the dopant's concentration. For every dopant concentration there is a characteristic pressure at which the electron concentration begins to decay with increasing oxygen pressure. This characteristic transition is also found in experiments for ITO [265] and has a temperature dependence such that it shifts to higher oxygen pressures at higher temperatures. Our selected temperature of 600° C corresponds to an average deposition/annealing temperature [138] and results for other temperatures can be obtained by shifting the oxygen pressure scale in the figure.

For the dopant concentration of 6×10^{20} cm^{-3} the decay of free electron concentration for SnO_2 is found at oxygen pressures which are more than eight orders of magnitude higher than those for In_2O_3. Further, for the lower dopant concentration of 2×10^{20} cm^{-3} neither of the materials suffers from compensation effects in the whole range of oxygen pressures.

Our findings are surprising in light of the experiments, since the conductivities reported for ITO are generally higher than in any TCO material related to SnO_2. Consequently, according to our hybrid-DFT results the origin of the significantly lower free electron concentrations in SnO_2 cannot be due to intrinsic acceptor defects. Instead, the results suggest that the limitations of n-type doping in SnO_2 mainly arise

due to dopant-specific properties rather than properties intrinsically related to SnO_2. A further optimization of *n*-type doping with respect to the resulting free electron concentrations is, therefore, possible.

The uncertainties, which are still connected to the actual experimental band-gap of In_2O_3, cannot alter our conclusions. Namely, by using the largest presently suggested band gap value for In_2O_3 [97, 95] for tuning the range-separation parameter of the XC-functional does not lead to increased formation energies for the Fermi level at the CBM. Band gap related ambiguities do not arise in the case of SnO_2. Note also that this result is not obtained using LDA. In this case both materials suffer from compensation effects at any dopant concentration and oxygen pressure considered. In addition, the doping limit is higher for In_2O_3 compared to SnO_2 using LDA. This qualitative change when going from LDA to hybrid functionals is due to the extraordinary large band gap error for SnO_2.

For FTO the doping limit is conjectured to be caused by interstitial fluorine defects [138, 266] which is now convincingly supported by our calculations. To the best of our knowledge the reason for the low electron concentrations of ATO is presently not known. Beside the occurrence of Sb^{3+} instead of Sb^{5+}[267, 268] which could act as acceptor, the segregation of the Sb-dopant is a possible, often-considered origin for low electron concentrations [269]. However, our calculations predict a low abundance of cation vacancies and, thereby, a good kinetic stability of cation dopants against vacancy-mediated migration in SnO_2. The kinetics will therefore be slow even if there is a tendency for segregation. More specifically our calculations explain that segregation cannot be observed for ATO [138] but it is possible for ITO [262] using a comparable experimental setup. As we have shown in Fig. 45, higher cation vacancy concentrations can be expected for doped In_2O_3.

8.3 SUMMARY AND CONCLUSION

We have reinvestigated the electron compensation in two TCO materials In_2O_3 and SnO_2. We have shown that within the hybrid-functional-DFT description In_2O_3 and SnO_2 are highly-*n*-type dopable against the formation of intrinsic acceptors. We have obtained a doping limit for In_2O_3 which is in good agreement with experiments. This reflects the robustness of the methodology used. Most importantly, we have obtained for SnO_2 a doping-limit, which is beyond the experimentally observed one. We conclude that for SnO_2 the lower measured electron concentrations are, therefore, not a consequence of any intrinsic acceptor of the material.

Our general result is that SnO_2 is more robust towards high Fermi-level values and should allow for a higher maximum doping than what has been reached in experiment so far. This conclusion is unlikely to be altered by any approximation used in our calculations. Our findings in turn state that the source for lower conductivities in SnO_2 in comparison with In_2O_3 are related to the dopants presently used (F and Sb). While the doping limit in In_2O_3 is given by intrinsic acceptors, the conductivities are

limited by other processes in SnO_2. Likely explanations for the presently observed doping limits are either the low ionization rate of the dopants, extrinsic acceptors, low solubility, or defect-defect interactions. Consequently, further improvements of cheap SnO_2 based TCO materials are possible by using other dopants and dopant combinations.

Part IV

POINT DEFECT KINETICS

In the following part we turn to the calculation of kinetic properties of point defects. This is necessary in order to estimate to what extent non-equilibrium concentrations of defects are possible. The important ingredient are the transition states for atomic jumps, which we obtain with the nudged elastic band method (NEB).

The first chapter deals with migration mechanisms of defects in In_2O_3. In this case, the complicated crystal structure leads to a wealth of different migrations paths and makes the evaluation of effective migration energies difficult. This problem is circumvented by the application of a kinetic Monte-Carlo method.

The second chapter then deals with the migration processes in SnO_2. The simpler rutile crystal structure allows the evaluation of finite size effects which is not possible in the case of In_2O_3.

Finally, we summarize important similarities and differences between SnO_2 and In_2O_3.

9

SELF-DIFFUSION IN INDIUM OXIDE

9.1 INTRODUCTION

Among the TCOs indium oxide (In_2O_3) and indium-tin oxide (ITO) play a particularly important role due to their excellent electronic properties in combination with a good processing characteristics and a high work function [264, 262], which makes them especially suitable for applications in devices containing organic functional layers [44].

Relevant point defects have been identified from measurements of electronic conductivity as function of oxygen partial pressure [118, 120, 109, 263] and from theoretical calculations within density-functional theory in previous chapters and by other first-principles studies [128, 129, 131, 130]. Surprisingly, rather little is known about the mobility of point defects in In_2O_3. In addition, there exist only a few studies which have attempted to measure diffusion coefficients in In_2O_3 [270, 271, 272]. In these studies it was not possible to reliably extract the migration energies of the migrating defects, since the defect concentrations were either not measured along with the total activation energy, [271, 272] or the specimen contained impurities [270].

In this study, we present a complete survey of self-diffusion processes of in In_2O_3. We systematically map the total energy surface by means of first-principles calculations for possible migration paths and identify the relevant processes for each defect type in several charge states. In particular, the role of the peculiar C-type rare earth structure (bixbyite) and its influence on the defect mobilities is addressed.

For consistency, we have adopted the methodology from Chapter 5, while the nudged elastic band (NEB) method is used for the barrier calculations (section 4.5.2). Further details about the computational setup can also be found in Ref. [A7]. The calculations were carried out in a supercell with 40 atoms and the number of images* used for the determination of energy barriers ranged between 3 and 7, depending on the complexity of the diffusion path. Effects of the supercell size were not evaluated due to the large unit cell. A discussion of cell size effects can be found for migration energies in SnO_2 in section 10.2.3.

*Initial and final state not included.

9.2 RESULTS AND DISCUSSION

9.2.1 *Migration mechanisms*

Various defect types can occur in In_2O_3 depending on the environmental conditions. The oxygen vacancies are important for the pure material [118, 119] and oxygen interstitials become important as soon as the material is *n*-type doped [109]. Defect studies have indicated the relevance of thermal indium vacancies [273, 258] as well as oxygen dumbbell defects at higher oxygen pressures [258, 131]. For completeness, we have also included the indium interstitials, although high formation energies were reported for this defect [258, 131]. The complete set of data including the energy barriers, jump lengths and jump directions, as well as further details about the migration processes are given in Tables 8, 9 and 10. In order to obtain the full set of migration barriers within the unit cell, all symmetry elements of the space group 206 ($Ia\bar{3}$) can be applied to the representant processes given in the tables. Anti-sites as well as the effect of association of defects are presently not considered. It was shown earlier, that formation energies of anti-sites are high while association energies are relatively small among donors (chapter 5). Further, we want to stress that in highly *n*-type doped material defect association and interaction with extrinsic impurities is certainly important [109, 112, 263], but not considered presently.

9.2.2 *Nomenclature*

In the following the different migration processes are labeled by the defect symbol according to standard notation in conjunction with a capital letter superscript indicating a distinct microscopic diffusional jump. The superscript may additionally contain a number to indicate that there is more than one path leading to the same final state (*e.g.* interstitialcy processes). In images the processes are indicated by arrows connecting two sites.

Furthermore, it is generally possible for the defects to occupy different sublattices (*e.g.* indium vacancies) sites. For this reason defect processes are also separated with respect to the sublattice of the initial state (compare with Tables 8, 9 and 10)

There is a total of seven symmetrically distinct sites (sublattices) to be considered for diffusion which are labeled with italic lower case letters according to Wyckoff notation (compare with Sect. 5.2 and 2.2.2). The general position (*e*) is relevant for oxygen vacancy diffusion while two sublattices (*b* and *d*) are involved for the cation vacancies. Both, oxygen and indium interstitials migrate on the structural vacancy site (*c*) whereas only for indium interstitials the high symmetry point (*a*) is a stable site. Interstitials have furthermore the possibility to migrate via indirect interstitialcy mechanism, which involves regular lattice sites being the *e*-site for oxygen interstitials and *b*/*d*-sites for the indium interstitials. Finally, the oxygen dumbbell defect occupies the regular oxygen lattice position (*e*) with two different atoms. This can be thought of as two different

general positions which are displaced from the original oxygen lattice site. In the following sections the lattice sites of the final states are also categorized in terms of neighbor shell relations and their respective crystallographic jump directions. For both, neighbor shells and directions the underlying fluorite lattice geometry is taken as reference in order to avoid large indices for directions and additional complications for neighbor relations.

9.2.3 *Vacancy mechanism*

Oxygen vacancy

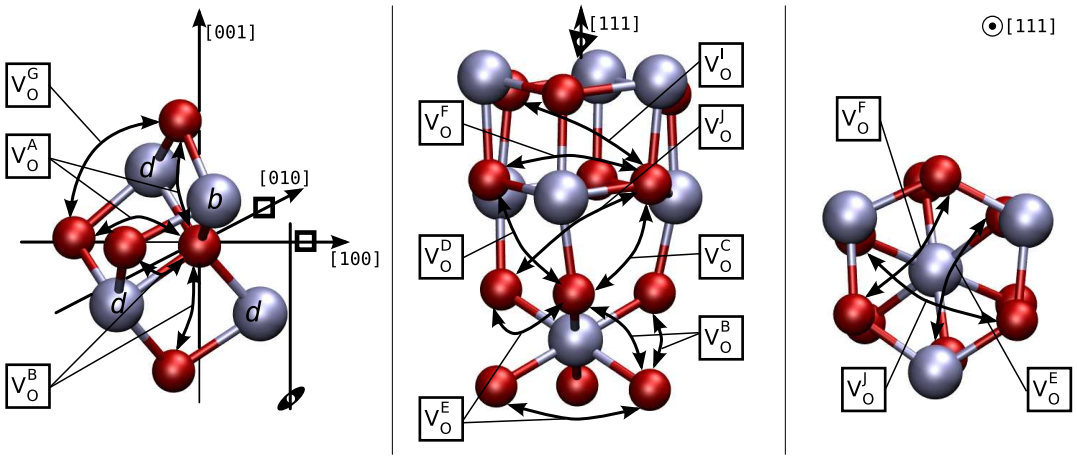


Figure 47: Nearest neighbors of the oxygen (left) and atomic arrangement surrounding the threefold axis (center / right). Rotation axis are included for orientation. Migration paths are indicated by their labels and the numeric values for the corresponding energy barriers can be found in Table 8.

The predominant charge state of V_O in In_2O_3 is still under debate [131, 274, 247, 275]. For this reason we have determined the energy barriers for the oxygen vacancies in all potentially occurring charge states ($q = 0,+1,+2$).

Fig. 47 shows the direct neighborhood of the oxygen vacancies. All nearest neighbor sites of the vacancy are located in $\langle 100 \rangle$ directions (Fig. 47 left). Migration processes to these target sites are only pairwise equivalent and denoted as processes $V_O^{\mathbf{A}}$ and $V_O^{\mathbf{B}}$, respectively. The two remaining $\langle 100 \rangle$ directions point towards structural vacancies (*c*-sites). Since these directions are not blocked by any atom, we have also considered migration over the structural vacancy, for which the target site is located at large distance within the fourth neighbor shell (process V_O^J in Fig. 47 center/right).

The second neighbor shell comprises 9 sites located in $\langle 110 \rangle$ directions from which 6 are geometrically accessible (not blocked by cations). They are connected by migration processes labeled $V_O^C - V_O^G$. The remaining three $\langle 110 \rangle$ directions again lead to vacant

positions. Processes $V_O^{\mathbf{D}}$, $V_O^{\mathbf{E}}$,$V_O^{\mathbf{F}}$ are twofold and represent mechanisms which lead to circulating motions with no net displacement, if considered separately (the same is true for processes $V_O^{\mathbf{A}}$ and $V_O^{\mathbf{B}}$). Transitions $V_O^{\mathbf{C}}$,$V_O^{\mathbf{G}}$ on the other hand are only onefold and their saddle points are aligned on the twofold rotation axis. Transition $V_O^{\mathbf{C}}$, appears to be of special interest, since its saddle point configuration has two structural vacancies in its proximity, and therefore exhibits the lowest oxygen vacancy barrier in any charge state (*e.g.* $\Delta E^M[+2] = 0.71$ eV). Note that this process represents a migration to the second neighbor shell in $\langle 011 \rangle$ direction (Fig. 47 center).

The six third neighbors are located in $\langle 111 \rangle$ directions among which three processes are completely blocked by cations and therefore are not considered as possible migration processes. Additionally, another two sites are vacant within this neighbor shell. Only $V_O^{\mathbf{H}}$ and $V_O^{\mathbf{I}}$ remain as possible migration processes, where $V_O^{\mathbf{I}}$ corresponds to a migration via an *a*-interstitial site. The corresponding migration energies are, however, rather high ($\Delta E^M[+2] \sim 5$ eV).

In total, we have calculated the barriers for 10 symmetrically distinct transitions (see Table 8), giving rise to 17 transitions in total. A remarkable feature of the oxygen vacancy diffusion is that each specific transition occurs with a maximum multiplicity of two. Apparently, not every site on the oxygen sublattice can be reached using only one specific type of process. At least two active mechanisms are necessary in order to allow for larger displacements within the oxygen sublattice. Our findings imply a mixed $V_O^{\mathbf{C}}$,$V_O^{\mathbf{A}}$ motion at low temperatures and some additional contribution of $V_O^{\mathbf{B}}$, $V_O^{\mathbf{E}}$ and $V_O^{\mathbf{D}}$ at elevated temperatures. Transitions via long jumps through the structural vacancies ($V_O^{\mathbf{J}}$) have a surprisingly low energy barrier ($\Delta E^M[+2] = 2.55$ eV) when compared with other processes of similar jump distance ($V_O^{\mathbf{H}}$, $V_O^{\mathbf{I}}$). This energy barrier is, however, still too high in order to significantly contribute to oxygen diffusion at relevant temperatures.

Indium vacancy

For In_2O_3, cation vacancy diffusion is assumed to be negligibly small [270]. This is partly because of their presumably low concentrations[109], however, not necessarily the case at higher temperatures [273] or in highly *n*-type doped samples, where the cation vacancy concentration can become more significant. Moreover, indium vacancies provide a possible migration path for the technologically important cation dopants (tin) for which a considerable mobility has already been reported in the literature [262].

The formation energies of vacancies of the two distinct indium lattice sites $V_{\text{In-b}}$ and $V_{\text{In-d}}$ differ by less than $\Delta E \sim 0.2$ eV for all possible charge states [131, 258]. Since there are two different sites with similar stability, both sublattices may contribute to material transport via indium vacancies. Therefore, it is useful to discuss individual jump processes which connect only *b*- or *d*-sites and mixed processes connecting *b*- with *d*-sites, separately. The predominant charge state over a wide range of Fermi energies was shown to be $q = -3$ on which we also focus here [131, 258]. We have,

however, also calculated the barriers for $q = -2$ in order to study the charge state dependence of migration.

There is only one process of the first type (*b* to *b*), which can be considered as a possible migration path. The target site is located towards $\langle 100 \rangle$ directions and separated by a rather large distance of 5.05 Å. There is no blocking along the migration path, but due to the large migration distance the energy barrier is high as $\sim$ 5 eV (process V_{In}^{C}) and therefore not activated at relevant temperatures.

Sites of *d*-type can be found in closer proximity, *i.e.* the nearest neighbors for In-*b* consist only of In-*d* which are distributed in a distorted face centered cubic arrangement around In-*b* (Fig. 48). Two symmetrically inequivalent migration processes can be identified within this nearest neighbor shell of In-*b*. Taking a (111) plane as the reference as it is denoted in Fig. 48 processes V_{In}^{A} are in-plane motions whereas V_{In}^{B} has out of plane components which lead to significantly different migration energies. While the barrier for the in-plane motion is 2.46 eV the out of plane migration is suppressed ($>$ 6 eV) due to close encounters with anions leaving process V_{In}^{A} the only relevant exchange mechanism between *b*- and *d*-sites.

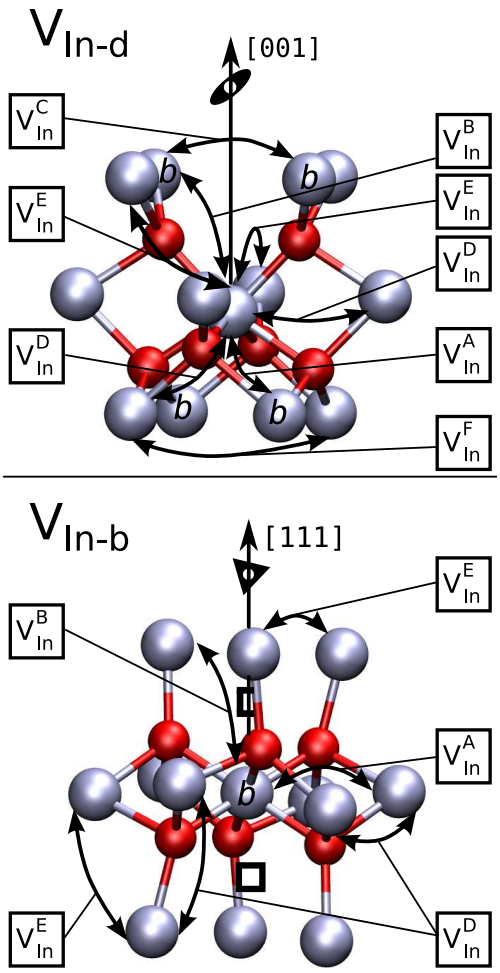


Figure 48: The nearest neighbors of the indium *d*- (top) and indium *d* -sites (bottom). Boxes indicate the locations of structural vacancies. Rotation axis are included for orientation. Migration paths are indicated by their labels and the numeric values for the corresponding energy barriers can be found in Table 8.

A greater diversity of transitions is found originating from *d* sites due to their lower symmetry. For the nearest neighbor shell four distinct processes can be identified (V_{In}^{A},V_{In}^{B}, V_{In}^{D} and V_{In}^{E}) instead of two (for In-*b*). Transitions V_{In}^{A} and V_{In}^{B} are both twofold and connect back to the In-*b* sites. The energy barriers differ by

the difference of defect formation energy between the indium vacancies on the two sites. The remaining processes $V_{In}^{\mathbf{D}}$ and $V_{In}^{\mathbf{E}}$ are both fourfold and lead to continued migration on the In-*d* sublattice with comparably low barriers of 2.42 eV and 2.08 eV, respectively. Processes at larger distances involve energy barriers higher than 4 eV and do not contribute to diffusion (processes $V_{In}^{\mathbf{F}}$, $V_{In}^{\mathbf{G}}$, and $V_{In}^{\mathbf{H}}$).

In total, the low energy paths involve the process $V_{In}^{\mathbf{A}}$ which connects the two sublattices *b* and *d* as well as transitions $V_{In}^{\mathbf{E}}$ connecting only *d*-sites. The *b*-site is by 0.2 eV more favorable when compared with *d*-sites. This results suggest that indium vacancies have a longer residence time on *b*-sites. However, once a vacancy moves, it will be located on a symmetrically different site, and migration takes mainly place on the other sublattice. The corresponding migration process on the In-*d* sublattice offers a complete set of transitions to reach any site within the lattice. The diffusion is mediated via *d*-sites although *b*-sites are more stable.

Defect	Process	λ [Å]	ΔE^M (q)[eV]							Z	r[Å]
			-3	-2	-1	0	$+1$	$+2$	$+3$		x, y, z
V_O	$V_O^{\mathbf{A}}$	2.79				1.73	1.09	0.99		2	(0.08, 0.44, -2.75)
$(x = 0.390)$	$V_O^{\mathbf{B}}$	2.81				2.00	1.37	1.28		2	(0.44, 0.36, 2.75)
$(y = 0.154)$	$V_O^{\mathbf{C}}$	2.94				1.53	0.77	0.71		1	(-2.22, -1.93, 0.00)
$(z = 0.382)$	$V_O^{\mathbf{D}}$	3.29				2.08	1.53	1.50		2	(0.44, -2.30, -2.30)
	$V_O^{\mathbf{E}}$	3.31				2.07	1.40	1.35		2	(0.08, -2.38, 2.30)
	$V_O^{\mathbf{F}}$	3.54				2.09	1.76	1.80		2	(-2.30, 2.67, 0.36)
	$V_O^{\mathbf{G}}$	3.71				3.93	–	3.82		1	(2.83, 0.00, -2.39)
	$V_O^{\mathbf{H}}$	4.47				–	–	5.72		2	(-2.67, -2.30, 2.75)
	$V_O^{\mathbf{I}}$	4.51				–	–	4.72		1	(-2.22, 3.11, -2.39)
	$V_O^{\mathbf{J}}$	4.62				2.65	–	2.55		2	(0.08, -4.61, 0.36)
$V_{\text{In-b}}$	$V_{\text{In}-\text{b}}^{\mathbf{A}}$	3.34	2.46	2.31						6	(0.00, -2.19, 2.52)
$(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$	$V_{\text{In}-\text{b}}^{\mathbf{B}}$	3.82	6.85	–						6	(0.00, -2.86, -2.52)
	$V_{\text{In}-\text{b}}^{\mathbf{C}}$	5.05	4.97	–						6	(0.00, 0.00, -5.05)
$V_{\text{In-d}}$	$V_{\text{In}-\text{d}}^{\mathbf{A}}$	3.34	2.27	2.29						2	(2.19, -2.52, 0.00)
$(x = 0.467)$	$V_{\text{In}-\text{d}}^{\mathbf{B}}$	3.82	6.65	–						2	(-2.86, -2.52, 0.00)
$(y = 0.0)$	$V_{\text{In}-\text{d}}^{\mathbf{D}}$	3.36	2.42	2.30						4	(-0.33, -2.52, 2.19)
$(z = \frac{1}{4})$	$V_{\text{In}-\text{d}}^{\mathbf{E}}$	3.83	2.08	2.11						4	(-0.33, -2.52, -2.86)
	$V_{\text{In}-\text{d}}^{\mathbf{F}}$	5.05	4.54	–						2	(-5.05, 0.00, 0.00)
	$V_{\text{In}-\text{d}}^{\mathbf{G}}$	5.10	4.54	–						2	(-0.67, 0.00, -5.05)
	$V_{\text{In}-\text{d}}^{\mathbf{H}}$	5.10	4.98	–						2	(0.00, -5.05, 0.00)

Table 8: Collection of migration barriers for bixbyite indium oxide including oxygen and indium vacancy mechanism. The first column contains the defect type and its location within the unit cell. The label used in images is given in column three followed by the corresponding jump distance λ and the migration energies ΔE^M in the relevant charge states. Column five gives the multiplicity (Z) of the corresponding process (number of times the process is available by symmetry from one site). The last column contains one representant of the corresponding jump vector (r).

9.2.4 *Interstitial mechanism*

Oxygen interstitial

In the bixbyite lattice the 16*c* positions offer a relatively large volume for the accommodation of interstitial defects. Indeed, the oxygen interstitials are of major technological interest for the *n*-type doped material [109, 112]. In the charge state $q = -2$ they are conjectured to be the main source of electron compensation in ITO [109]. However, the unit cell comprises 80 atoms and 16 interstitial sites are therefore well separated from each other. The neighbor relations of the interstitial positions are depicted in Fig. 49 and the corresponding energy barriers can be found in Table 9. The interstitial sites are aligned chainlike in $\langle 111 \rangle$ directions and alternatingly separated by interstitial *a*-positions and In-*b* cations (Fig. 23).

The closest neighbor is reached just along the $\langle 111 \rangle$ direction. This is a one-fold process since the opposite direction ($[\bar{1}\bar{1}\bar{1}]$) is blocked by an In-*b* cation. This process ($O_i^{\mathbf{A}}$) involves traversing an interstitial *a*-site, which for oxygen interstitials in charge state $q = -2$ is a saddle point configuration.

The three second neighbors are located towards $\langle 110 \rangle$ directions and perpendicular to process $O_i^{\mathbf{A}}$. The jump distance of this process $O_i^{\mathbf{B}}$ is rather large (3.63 Å) and additionally cations partly obstruct the diffusion path (Fig. 49). Therefore, the corresponding interstitialcy processes involving the regular lattice sites (processes $O_i^{\mathbf{B1}}$ and $O_i^{\mathbf{B2}}$ in Fig. 49) need to be considered and compared with the direct process. The interstitialcy processes may occur via two different lattice oxygen sites depicted in Fig. 49. By symmetry, however, both processes $O_i^{\mathbf{B1}}$ and $O_i^{\mathbf{B2}}$ are equivalent in energy and offer an effective multiplicity of six compared to three for the direct process. The interstitialcy processes have significantly lower migration energies when compared with the direct ones (1.38 eV vs. 2.42 eV). Furthermore, interstitialcy jumps to even larger distances can be constructed (process $O_i^{\mathbf{D1}}$/ $O_i^{\mathbf{D2}}$) with only slightly higher energy barriers (1.48 eV). In this cases, no direct transition is geometrically realizable. Similarly, the path $O_i^{\mathbf{E}}$ in $[\bar{1}\bar{1}\bar{1}]$ direction is blocked by an indium atom (Fig. 49) but two interstitialcy processes ($O_i^{\mathbf{E1}}$, $O_i^{\mathbf{E2}}$) can be constructed, which are again symmetrically equivalent. These events have high migration energies and are not likely to play a role.

The lowest migration barrier for oxygen interstitial diffusion is process $O_i^{\mathbf{A}}$ (1.22 eV). Similar like in the case of oxygen vacancies this single process is not sufficient to provide a complete diffusion path. In order to reach any interstitial site of the lattice, interstitialcy jumps via $O_i^{\mathbf{B1}}/O_i^{\mathbf{B2}}$ are needed. The rate determining transitions are therefore $O_i^{\mathbf{B1}}$ and $O_i^{\mathbf{B2}}$. In between these events $O_i^{\mathbf{A}}$ will be active, however, not lead to any net displacement but only to an exchange of the interstitial between two *c*-sites in nearest neighborhood.

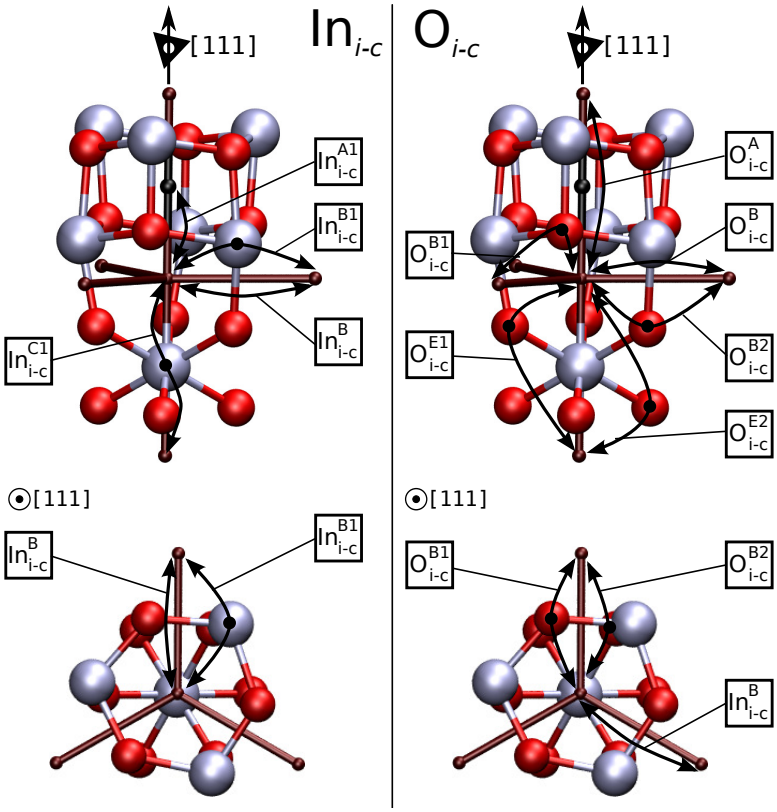


Figure 49: Nearest neighborhood of the interstitial positions for $In_{i\text{-}c}$ (left) and $O_{i\text{-}c}$ (right). The nearest interstitial positions are connected. The additionally stable stationary point (*a*-site) for indium interstitials is denoted by a black ball. Rotation axis are included for orientation. Migration paths are indicated by their labels and the numeric values for the corresponding energy barriers can be found in Table 9.

Indium interstitial

For indium interstitial diffusion the same sublattice as for the oxygen interstitials is used. The minimum energy position of the interstitial is located closer to the In-*b* atom when compared with the oxygen interstitials (compare values of first column in Table 9) Additionally, the interstitial *a*-site is a stable minimum with an increased formation energy of 0.6 eV for indium interstitials, whereas it is simply a saddle-point configuration for oxygen interstitials.

There is one unique process ($In_i^{\mathbf{A1}}$) connecting *c*- with *a*-positions. The decay back to a regular interstitial site can occur to either side along $\langle 111 \rangle$ (Fig. 49 left). In contrast to the case of oxygen interstitials, the $[\bar{1}\bar{1}\bar{1}]$ direction is available for migration via an interstitialcy mechanism ($In_i^{\mathbf{C1}}$). Energetically, the latter process also constitutes the lowest energy transition for indium interstitials with a remarkably low barrier of

0.42 eV in the charge state +3. The opposite direction towards the *a*-site is considerably higher (1.62 eV) but still lower in transition energy than processes leading away from the ⟨111⟩ axis (1.69 eV)†. Interstitialcy processes leading away from the ⟨111⟩ axis ($In_i^{\mathbf{B1}}$) are again favored over direct mechanisms ($In_i^{\mathbf{B}}$). Compared to oxygen interstitials the multiplicity of this mechanism is only three due to the smaller number of neighboring indium cations.

For indium interstitials comparable energy barriers were found as for oxygen interstitial diffusion ($\sim$ 1.3 eV-1.5 eV), however, the overall mechanism is very different. In the case of oxygen, transitions leading away from ⟨111⟩-axis ultimately determine the activation for diffusion although processes along ⟨111⟩ are always faster. For indium interstitials, the *b*-lattice indium does not play a constraining role within ⟨111⟩ direction and the data suggests another mode of migration. Here, the ⟨111⟩ axes can serve as quasi one dimensional diffusion channels with a sequence of processes $\ldots In_{i-c}^{\mathbf{C1}} \rightarrow In_{i-c}^{\mathbf{A1}} \rightarrow In_{i-a}^{\mathbf{A1}} \ldots$. Depending on temperature the interstitials can change to other ⟨111⟩-channels via $In_{i-c}^{\mathbf{B1}}$.

Oxygen dumbbell interstitial

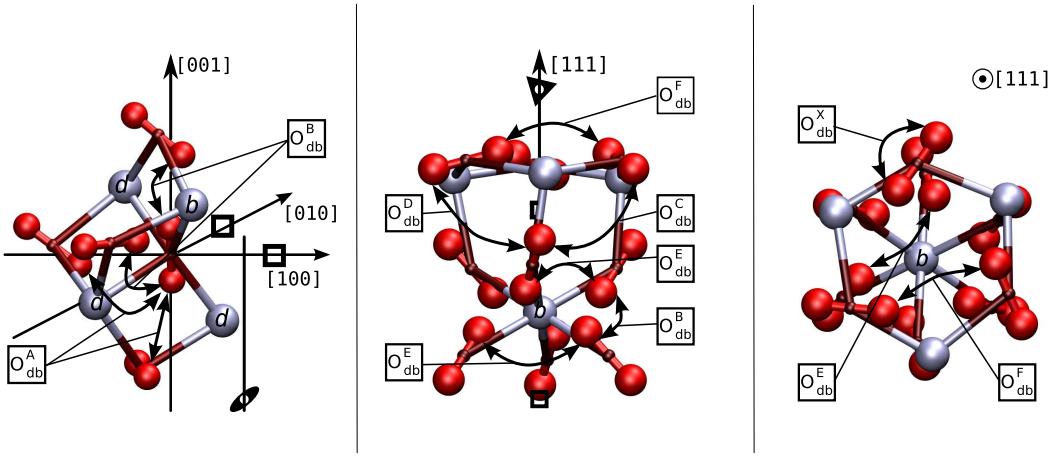


Figure 50: First neighbors of the oxygen site when occupied by a dumbbell defect (left). View along an ⟨111⟩ axis from the side (center) and from the top (right). The most important jump processes are labeled within the images.

The oxygen dumbbell interstitial is a covalently bonded oxygen dimer defect located on a regular lattice site. This defect can be found in several orientations (chapter 5). The most stable form of the dumbbell is the neutral charge state. In addition, we have shown that for the neutral charge state the regular oxygen interstitial *c*-sites

†Although the energy barriers do not differ much we have found indications that, the difference would increase for larger cell sizes.

denote only very shallow or no minima on the total energy surface and can therefore be considered as transition states for the oxygen dumbbell rather than stationary intermediate configurations [A5]. Here, only the dumbbell configuration in its most stable configuration is considered for migration.

Since the motion of the dumbbell is exclusively restricted to the regular oxygen sublattice similar considerations apply for possible migrations mechanisms as for the oxygen vacancies. The difference, however, is that there are now two particles attempting a jump. As a consequence, each process of the oxygen vacancy diffusion generally results in four distinct processes. For the same reason also Table 10 is split into two parts for $O_{i,db-1}$ and $O_{i,db-2}$, respectively. The $O_{i,db-1}$ is the oxygen, which has In-*b* as nearest cation neighbor (Fig. 50 left).

In Fig. 50 the neighboring oxygen dumbbell orientations of the first and second neighbor shells of the oxygen sites are shown. For each oxygen site two atoms are depicted with their varying relative dimer axis orientations. Due to the vast number of possible processes we have only considered transitions for which the analogous oxygen vacancy mechanism barriers did not exceed 2.5 eV. Moreover, the number of processes can further be decreased by symmetry and geometrical considerations. For example, we have mainly considered processes where the ground-state configurations of the dimer on initial and target sites assume an approximately head to head configuration.

Beside site to site jumps, the oxygen dumbbell can perform an on-site rotation *i.e.* instantaneous exchange of $O_{i,db-1}$ with $O_{i,db-2}$ atoms with an energy barrier of 0.87 eV.

The lowest corresponding migration energies for the dumbbell migration do neither comprise the first neighbor shell ($O^{\mathbf{A}}_{db-2}$and $O^{\mathbf{B}}_{db-1}$) nor process $O^{\mathbf{C}}_{db-1}$, for which the lowest migration energies were found in the case of the oxygen vacancies. These processes have migration energies of 1.40 eV, 1.68 eV and 2.06 eV for $O^{\mathbf{A}}_{db-2}$, $O^{\mathbf{B}}_{db-1}$ and $O^{\mathbf{C}}_{db-1}$, respectively. In the case of the oxygen dumbbell, the lowest migration energies are found among jumps which direct the defect around the three fold rotation axis ($O^{\mathbf{E}}_{i,db-1}$ and $O^{\mathbf{F}}_{db-2}$) and are part of the second neighbor shell of the oxygen sublattice. This is plausible since the dumbbell consists of a surplus atom and the structural vacancies are all aligned on three-fold axis. Fig. 50 shows the relative orientations of dimers in an arrangement close to the three fold rotation axis. The exchange process for these transitions is probably facilitated by the presence of the structural vacancy which is not the case for the nearest neighbor processes ($O^{\mathbf{A}}_{db-2}$ and $O^{\mathbf{B}}_{db-2}$). These two low energy processes, however, do not suffice to provide migration to all oxygen lattice sites. For a net displacement at least three different migration mechanisms need to be active. Our results suggest that beside the low energy processes $O^{\mathbf{E}}_{i,db-1}$ (1.01 eV) and $O^{\mathbf{F}}_{db-2}$ (1.20 eV) process $O^{\mathbf{D}}_{db-1}$(1.35 eV) needs to be additionally active (see Fig. 50).

Defect	Process	λ [Å]	ΔE^M(q) [eV]							Z	r[Å]	$\alpha \rightarrow \lambda_2$ [Å]
			-3	-2	-1	0	$+1$	$+2$	$+3$		x, y, z	
O_i	$O_i^{\mathbf{A}}$	3.62		1.22	0.52					1	(2.09, 2.09, 2.09)	
($x = 0.397$)	$O_i^{\mathbf{B}}$	3.63		2.42	–					3	(2.09, -2.96, 0.00)	
($y = 0.397$)	$O_i^{\mathbf{B1}}$	2.22		1.38	0.77					3	(-0.15, 2.15, -0.51)	$101.6 \rightarrow 2.46$
($z = 0.397$)	$O_i^{\mathbf{B2}}$	2.45		1.38	0.77					3	(-0.06, -2.45, -0.15)	$101.6 \rightarrow 2.22$
	$O_i^{\mathbf{C1}}$	2.45		2.98	–					3	(-0.06, -2.45, -0.15)	$133.1 \rightarrow 3.90$
	$O_i^{\mathbf{C2}}$	3.90		2.98	–					3	(-0.15, -2.90, 2.60)	$133.1 \rightarrow 2.46$
	$O_i^{\mathbf{D1}}$	2.45		1.47	–					3	(-0.06, -2.45, -0.15)	$135.7 \rightarrow 3.44$
	$O_i^{\mathbf{D2}}$	3.43		1.47	–					3	(-0.06, 2.60, 2.24)	$135.7 \rightarrow 2.46$
	$O_i^{\mathbf{E1}}$	2.45		4.16	–					3	(-0.06, -2.45, -0.15)	$100.7 \rightarrow 4.08$
	$O_i^{\mathbf{E2}}$	4.07		4.16	–					3	(-0.51, -2.81, -2.90)	$100.7 \rightarrow 2.46$
$In_{i\text{-}c}$	$In_{i-c}^{\mathbf{A1}}$	2.23						1.36	1.62	1	(1.29, 1.29, 1.29)	
($x = 0.372$)	$In_{i-c}^{\mathbf{B}}$	3.57						2.16	2.75	3	(0.00, 2.58, -2.47)	
($y = 0.372$)	$In_{i-c}^{\mathbf{B1}}$	2.42						1.64	1.69	3	(-1.23, 1.62, 1.29)	$95.35 \rightarrow 2.42$
($z = 0.372$)	$In_{i-c}^{\mathbf{C1}}$	2.14						0.29	0.42	1	(-1.23, -1.23, -1.23)	$180.00 \rightarrow 2.15$
$In_{i\text{-}a}$	$In_{i-a}^{\mathbf{A1}}$	2.23						0.67	0.99	2	(-1.29, -1.29, -1.29)	

Table 9: Collection of migration barriers for oxygen and indium interstitials. The first column contains the defect type and its location within the unit cell. The label used in images is given in column three followed by the corresponding jump distance λ and the migration energies ΔE^M in the relevant charge states. Column five gives the multiplicity (Z) of the corresponding process (number of times the process is available by symmetry from one site). Column six (r) is one representant of the corresponding jump vector. The last column contains the angle (α) and jump distance (λ_2) of the lattice atom in the case of an intersticialcy mechanism.

Defect	Process	λ [Å]	$\Delta E^M(q)$[eV]							Z	r[Å]	Sites
			-3	-2	-1	0	$+1$	$+2$	$+3$		x, y, z	
$O_{i,db-1}$	O^{X}_{db-1}	1.52				0.87				1	(0.76, -0.14, 1.30)	$O_{i,db-1} \rightarrow O_{i,db-2}$
($x = 0.390$)	O^{Y}_{db-1}					2.46				2		$O_{i,db-1} \rightarrow O_{i-a}$
($y = 0.154$)	O^{Z}_{db-1}					1.43				1		$O_{i,db-1} \rightarrow O_{i-c}$
($z = 0.382$)	O^{B}_{db-1}	1.74				1.68				2	(-0.13, 0.21,-1.71)	$O_{i,db-1} \rightarrow O_{i,db-1}$
	O^{C}_{db-1}	3.22				2.06				1	(2.22, 1.93, 1.30)	$O_{i,db-1} \rightarrow O_{i,db-2}$
	O^{D}_{db-1}	2.82				1.35				1	(2.03, -0.90,-1.72)	$O_{i,db-1} \rightarrow O_{i,db-2}$
	$O^{E}_{i,db-1}$	2.52				1.00				2	(-0.35, 1.93,-1.58)	$O_{i,db-1} \rightarrow O_{i,db-1}$
$O_{i,db-2}$	O^{X}_{db-2}	1.52				0.87				1	(-0.76, 0.14, -1.30)	$O_{i,db-2} \rightarrow O_{i,db-1}$
($x = 0.390$)	O^{A}_{db-2}	2.18				1.40				1	(0.18, -0.76, 2.03)	$O_{i,db-2} \rightarrow O_{i,db-2}$
($y = 0.154$)	O^{C}_{db-2}	3.22				2.06				1	(2.22, 1.93, -1.30)	$O_{i,db-2} \rightarrow O_{i,db-1}$
($z = 0.382$)	O^{D}_{db-2}	2.82				1.35				1	(-0.90, 1.72, 2.03)	$O_{i,db-2} \rightarrow O_{i,db-1}$
	O^{F}_{db-2}	2.73				1.19				2	(1.27, -2.21, -0.94)	$O_{i,db-2} \rightarrow O_{i,db-2}$

Table 10: Collection of migration barriers for oxygen dumbbell defect. The first column contains the defect type and its location within the unit cell. The labels of the processes are given in column three followed by the corresponding jump distance λ and the migration energies ΔE^M in the relevant charge states. Column five gives the multiplicity (Z) of the corresponding process (number of times the process is available by symmetry from one site). Column six is one representant of the corresponding jump vector (r). The last column contains the sites which are connected by the respective process.

9.2.5 *Effective values and annealing temperatures*

In order to use our data for obtaining macroscopic diffusion parameters, the set of migration barriers for each defect in a certain charge state needs to be averaged in an appropriate way. From the findings in sections 9.2.3 and 9.2.4 it is clear that for all defects (except V_{In}) the lowest energy processes do not build a closed diffusion path. As it was also shown for several defects more than one sublattice is necessarily used during migration in In_2O_3. Consequently, the effective migration energy (which can be experimentally measured) will not be dominated by the migration processes with the lowest energy.

In order to appropriately evaluate the contribution of each barrier to the effective migration energy ΔE_M^{eff} we have adopted a kinetic Monte-Carlo algorithm [208] in order to evaluate the mean square displacement for each defect in its different charge states. We have performed random walks on the respective sublattices and included all calculated barriers. The mean square displacement was averaged for each temperature until ΔE_M^{eff} and λ_M^{eff}, the effective migration barrier and jump length, were converged to a relative error better than 10^{-2}. The resulting values are listed in Table 11 and given for two different temperature ranges. We find that, generally, both the effective jump length and the effective exponential factor become slightly temperature dependent parameters. For an estimate of the absolute jump frequencies, from the transition probabilities, effective attempt frequencies are additionally needed. The values can be approximated on the basis of phonon calculations, which we have presented previously [273] and which also agree with IR-absorption data [276]. For diffusion of oxygen species (V_O, O_i) and indium species (V_{In}, In_i) we have used 15 THz and 5 THz, respectively. They are obtained by averaging over the corresponding partial phonon densities of states and mainly account for the differing masses of the two constituents. The total effective jump frequency was also converted into an "annealing temperature", which indicates the temperatures at which a particular defect becomes significantly mobile *i.e.* when a specific defect type can equilibrate its population in the material in some reasonable time. As a definition we use the temperature at which the diffusion constant for the specific defect reaches a value of $10^{-12}\,\frac{cm^2}{s}$ and the corresponding temperature for each defect is listed in Table 11.‡ The value is largely dominated by the exponential term and does not severely suffer by the approximation of the attempt frequencies. Note, that ΔE_M^{eff} generally differs from the lowest migration barrier for any defect type which is also given in the same table.

‡This estimate is based on the Einstein equation and assumes the equilibration of a micrometer sized sample within an hour.

Defect	q	ΔE^M_{eff} [eV]	λ_{eff} [Å]	ΔE^M_{eff} [eV]	λ_{eff} [Å]	ΔE^M_{min} [eV]	λ_{min} [Å]	T_{an} [$^\circ K$]
		1700K - 1500K		700K - 900K				
V_O	0	1.83	3.49	1.77	2.44	1.53	2.94	640
	+1	1.20	3.58	1.13	2.39	0.77	2.94	410
	+2	1.10	3.32	1.02	2.26	0.71	2.94	370
V_{In}	-3	2.26	4.09	2.27	4.24	2.21	3.83	820
	-2	2.21	4.13	2.17	3.35	2.11	3.83	800
O_i	-2	1.41	6.36	1.40	5.96	1.22	3.62	480
	-1	0.74	3.94	0.77	4.62	0.52	3.62	270
$O_{i,db}$	0	1.19	2.52	1.20	2.65	1.00	2.52	430
In_i	+3	1.67	7.27	1.66	7.08	0.42	2.14	580
	+2	1.44	5.99	1.39	4.41	0.29	2.14	500

Table 11: Effective migration energies ΔE^M_{eff} and effective diffusion length λ for all defects in two different temperature regimes. Additionally, the lowest energy path with its corresponding jump distance is given. The effective migration energy is generally higher when compared with the lowest migration energies. The last column contains the annealing temperature which indicates the temperature range in which the defect becomes significantly mobile.

Mobile defects

From Table 11 it can be readily seen that oxygen vacancies in their presumably most stable charge state +2 become significantly mobile already at slightly elevated temperatures of about 400K. In contrast, the indium vacancies do not get mobile before reaching temperatures of at least 820K, which is not relevant in most technical applications but for diffusion measurements [271, 270]. The difference is much less pronounced for the interstitials in the predominant charge states +3 and −2 for indium interstitials and oxygen interstitials, respectively. They become mobile in a temperature range around 500K-600K, which are relatively high values for interstitial diffusion. The dumbbells on the other hand have an annealing temperature comparable with that of the oxygen vacancies in charge state +2. To our knowledge, no dedicated annealing experiments (after irradiation) have been performed on In_2O_3 yet, so that comparison can only be done indirectly. Wirtz and Takiar [270] have derived the chemical diffusion constant from relaxation measurements and extracted 1.5 eV and 1.4 eV for the migration energies of $O_{i\text{-}c}^{2-}$ and V_O^{2+}, respectively. Considering the statistical accuracy of their measurement and the impure sample, the agreement with our calculations (1.4 eV and 1.10 eV) reasonable and within the error bounds. It is especially noteworthy that the oxygen interstitial migration is found to have a larger effective migration energy in both, theory and experiment. Furthermore, studies on the electrical conductivity[118, 120, 119] show annealing effects (irreversible change of conductivities) at temperatures of $\sim$ 400K which can be related to the activation of oxygen vacancy migration in charge state +2.

Role of the structural vacancy

The presence of the vacant sites in the bixbyite structure influences diffusion in two different ways. Firstly, the point symmetry of all sublattices is lowered and generally a greater variety of different barrier heights can be identified. Due to the increased free volume, jumps towards this interstitial sites are usually lowered. However, for the oxygen and interstitial sublattices, the lowest barriers do not offer a complete migration path via which a net diffusion can occur for a defect as we have shown. The lowest energy barriers therefore do not contribute to the migration energy in these cases, but to the prefactor for diffusion by connecting two lattice sites and increasing the multiplicity of the remaining mechanisms. This also means that the available free volume in In_2O_3 when compared with other CaF_2 derived structures does not result in a necessarily higher diffusion in the material as far as the barriers are concerned. Nevertheless, the structural vacancies enable the structure to accommodate large numbers of interstitial defects, which contribute to the tracer diffusion coefficient via the defect concentrations. The main effect of the structural vacancies is the break in symmetry and the resulting break of degeneracy in the migration processes.

Charge state dependence

As it is shown in Fig. 51 a charge state dependence of migration energies for all defects is found. The effect is mostly homogeneous for all processes (oxygen interstitials are an exception) of one particular defect type. While the migration energies tend to be low in the case of lower charge states, the opposite is true for the oxygen vacancies. Generally, the migration energy for V_O is increased by ~ 0.7 eV when compared with the ionized vacancies ($q = +1/+2$). This can be explained by the distance dependent relaxation behavior of the vacancies described in chapter 6. For migration, the atomic arrangement and distances around the vacancy state are changed and the vacancy state moves up within the band gap. In the calculation, this is indicated by a heavy occupation of the CB. The artificially high Fermi level leads to an excess energy in the barrier calculations. Since the occupation of the CB is also depending on cell size a major part of the difference in migration energies is attributed to a finite size effect, which was, however, not further investigated.

In the case of acceptor defects the charge state dependence of migration energies is most important for oxygen interstitials and to a lesser extent for indium vacancies. Since the $q = -2$ and $q = -3$ charge states are most stable for these defects, the charge state dependence has another implication. For these charge states electrons occupy acceptor states (Fig. 25) close to the VBM.

The absorption of light with energies little below the band gap could remove electrons from these states. The thus obtained deionized defects (lower charge state) have a higher mobility and could lead to a light induced enhancement of diffusion.

The lower migration energies can easily be understood considering the structural similarity between the saddle point configuration for intersticialcy migration (process $O_i^{\mathbf{B1}}$,$O_i^{\mathbf{B2}}$) and the ground state geometry of $O_{i,db}$. Electron removal from interstitial oxygen results in a O_i^- radical, which is unstable and, therefore, chemically reactive. The created electron hole can be compensated by recapturing an electron or by ejecting a second electron from the oxygen and from $O_{i,db}$. The latter scenario would bring the migrating oxygen into the saddle point configuration.

In experiments light induced enhancement of equilibration has already been observed [277] and is now well explained by our findings. We suggest the oxygen interstitials to be responsible for enhanced oxygen diffusion under UV illumination. Furthermore, we point out that this effect could also lead to aging effects in optoelectronic devices especially with organic materials and with short wavelengths of light due to an enhanced leakage of oxygen from ITO electrodes.

9.3 ACCURACY

The numerical accuracy of the total energy calculations does only marginally affect the total error of the defect calculations, when compared with other potential sources of errors. Energy differences are taken of very similar configurations with the same number of atoms so that numerical inaccuracies should greatly cancel.

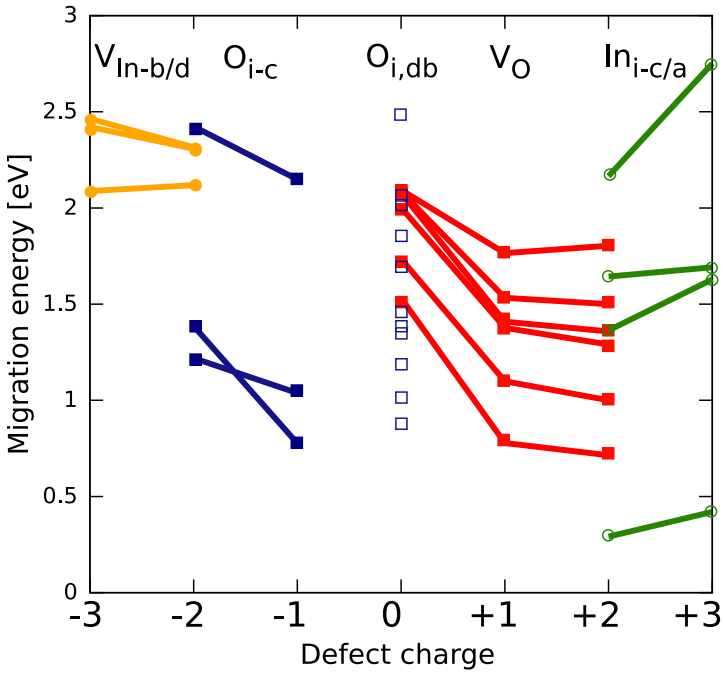


Figure 51: The charge state dependence for several barriers and all defect types. The order of barrier heights is mostly kept with changing charge state.

We assume the errors of the calculation to arise either from finite cell size effects or due to problems connected with the XC-functional. Using the GGA functional in conjunction with the on-site corrections, we obtain a realistic picture of the bonding behavior and an accurate cohesive energy and cell volume. In particular we found that the relaxation energies of defects improved using this method (chapter 6). We therefore assume the GGA+U method to be superior to GGA only, because it reproduces best the electronic structure as well as the structural parameters at low computational costs. Tests, however, showed that the differences in energy barriers are in the range of ~ 0.1 eV when compared with GGA calculations.

Besides the XC-functional, the major errors can be expected to emerge from the finite cell size. Since the Madelung-type electrostatic finite-size contributions cancel for defect migration, the energies are expected to converge more rapidly and scale like $1/N$ or faster (elastic interactions), where N is the number of atoms. Due to the large number barriers, the complicated crystal structure and the fact that several different cell sizes are needed in order to evaluate the scaling behavior we have not investigated the cell size dependence, presently. Assuming that the cell size effect is

mainly caused by elastic strain fields the calculated values should be seen as an upper boundary for the migration energies in the diluted limit. We assume that the relative change of migration energies would be small when compared with the errors on an absolute scale. Therefore, the qualitative conclusions should not be affected by finite size effects. It is also noteworthy that semi-local functionals tend to underestimate the energy barriers,[278, 279] which may lead to fortious cancellations in our study.

9.4 SUMMARY AND CONCLUSION

For the oxygen vacancy and interstitials, we have shown that the break of symmetry and the omission of anion sites compared with the fluorite lattice introduces migration barriers with considerably low energies. However, these barriers can only contribute to the prefactor at relevant temperatures. The oxygen vacancy diffusion is mediated by two of the four nearest oxygen neighbors plus one process to the second neighborhood. For the interstitials we have found diffusion to occur via a mixed interstitial/interstitialcy mechanisms. The regular interstitial mechanism connects interstitial sites along $\langle 111 \rangle$ axis whereas jumps away from the $\langle 111 \rangle$ axis need to proceed via an interstitialcy mechanism for indium as well as oxygen interstitials, where the multiplicity is higher for oxygen interstitials.

Migration is mediated by processes along $\langle 111 \rangle$ for the indium interstitials, whereas the perpendicular direction is rate determining for the oxygen interstitials. Indium vacancy diffusion is found to have energy barriers of about 2 eV. The diffusion mechanisms are such that the indium sites with higher symmetry are avoided most of the time. The lowest migration barrier can provide a complete path throughout the structure. Lastly, we have investigated oxygen dumbbell diffusion which occurs on the oxygen sublattice. We found that the normal interstitial positions are avoided in this case. The nearest neighbors do not significantly contribute to diffusion of this defect. Dumbbell defects mainly perform on-site rotations and dumbbell exchange takes place along $\langle 110 \rangle$ directions within $\{111\}$ planes which are part of the second neighbor shell. The symmetry is strongly broken and there is a whole ladder of possible transition energies. A strong charge state dependence is found for migrating oxygen rather than for indium atoms and is systematic for different migration processes.

We have presented a comprehensive study on diffusion mechanisms in indium oxide of the cubic polytype. To our knowledge it is the first attempt to quantify the kinetics of point defect migration in a C-type rare earth crystal structure (bixbyite) by means of first principles calculations. We have mapped out the total energy surface in In_2O_3 for possible migration paths for several defects in their predominant charge states. We find a wealth of contributing migration processes, especially for defects which are bound to the oxygen sublattice (V_O, $O_{i,db}$). For most of the defects, we have shown that the low energy processes do not dominate the effective exponential factor, which is the measured quantity in experiments. We have consistently determined the effective parameters for diffusion by means of KMC simulations. The oxygen

interstitial migration was found to have higher migration energy in accordance with experiments and a migration which is fundamentally distinct from that of the cation interstitials. The latter defects migrate quasi-one dimensionally along $\langle 111 \rangle$ axis, whereas for oxygen interstitial ($q = -2$) migration the determining processes are jumps perpendicular to the $\langle 111 \rangle$ axis. The indium vacancies have been found to migrate slower, but a measurable mobility should be expected for experiments at higher temperatures. The calculated annealing temperature of V_O^{2+} could be related to irreversible features in electrical conductivity measurements. Finally, the diffusion of oxygen interstitials is promoted by light irradiation with energies close to the band gap excitation.

10

SELF-DIFFUSION IN TIN OXIDE

10.1 INTRODUCTION

Since the observation of the gas sensing properties of SnO_2, a huge body of work has been devoted in order to obtain a detailed understanding of the gas sensing mechanism, surface structure and chemistry of the material [51] and to improve the selectivity of this material for different environments and gases [280, 45]. It was pointed out by Göpel that in the field of gas sensing properties most studies followed an empirical "trial and error" approach [281], while an in depth understanding of the microscopic processes is indispensable for further improvements of the material properties.

For the technologically important Taguchi sensors [48, 45], previous studies have mainly focussed on the surface and near surface areas of SnO_2, because the sensing mechanism was attributed to surface processes, only [51, 45, 282]. It is clear, however, that oxygen diffusion, exchange with the environment, adsorption and incorporation can also have a tremendous influence on the performance of such sensing devices. Although a number of experimental studies have investigated the bulk kinetics in SnO_2 [114, 115, 283, 284], to our knowledge the only studies of diffusion in single crystalline SnO_2 samples with controlled oxygen partial pressure were carried out by Kamp *et al.* [285, 255]. Their data will later serve as a benchmark for our theoretical results.

The aim of the present study is twofold: First, we conduct a survey on the possible migration paths for all defects in SnO_2 considering also different charge states. Second, we analyze the effect of supercell size in order to obtain reliable migration energies in the dilute limit. This procedure was not possible in the case of In_2O_3 due to its structural complexity (chapter 9). It is, however, assumed that insights form an analysis of finite cell size effects in SnO_2 can be transfered to In_2O_3. The finite size effects were corrected by using several supercell sizes of 72, 162, 216 and 270 atoms. The DFT setup is based on LDA and was adopted from chapters 7 and 8. Additional information is also given in Refs. [A8] and [A14].

The LDA/GGA+U method has been abandoned in the case of SnO_2 because the interaction of the VB with the semi-core 4d-states is negligibly weak (chapter 6). Removing the 4d-states speeded up calculations and enabeled us to conduct the finite

size scaling. We carried out additional barrier calculations using GGA-PBE [158] and PBE0 [244, 286, 287, 288, 166] to check the robustness of the selected method. For the barriers we used the NEB method again with the setup described in Chapter 9.

10.2 RESULTS

In the following we describe the migration processes for each individual point defect, separately. The different migration processes are labeled with the respective defect symbol with capital letters denoting the respective process. The values calculated within a 72 atomic supercell are presented in Table 12 from which we extract the rate controlling processes. In section 10.2.3 we refine the calculated energy barriers by subjecting them to a finite size scaling in order to check the sensitivity of our results with respect to the supercell size and to obtain values of the energy barriers for the dilute limit.

Table 12: Migration energies of the oxygen and tin vacancies. The data for the oxygen vacancy is including the two charge states. The directions are given in order indicate the isotropy of the migration process. The symbol λ is the jump distance of the respective process, while the multiplicity indicates the number of symmetrically equivalent processes of the same type. The distances λ are based on the LDA lattice constants $a = 4.731$ Å, $c = 3.198$ Å.

Defect	Process	Jump direction	λ[Å]	Mult.	E^M [eV]
V_O ($q = 0$)	$V_O^{\mathbf{A}}$	$\langle 110 \rangle$	2.59	1	2.58
	$V_O^{\mathbf{B}}$	$\langle 001 \rangle$	3.20	2	2.77
	$V_O^{\mathbf{C}}$	$\langle 111 \rangle$	2.90	8	2.09
	$V_O^{\mathbf{D}}$	$\langle 110 \rangle$	3.43	2	3.44
V_O ($q = +2$)	$V_O^{\mathbf{A}}$	$\langle 110 \rangle$	2.59	1	1.32
	$V_O^{\mathbf{B}}$	$\langle 001 \rangle$	3.20	2	2.42
	$V_O^{\mathbf{C}}$	$\langle 111 \rangle$	2.90	8	1.33
	$V_O^{\mathbf{D}}$	$\langle 110 \rangle$	3.43	2	3.13
$[V_O - V_O]$ ($q = +2$)	$V_O^{\mathbf{C}}$	$\langle 111 \rangle$	2.90	8	2.01
$O_{i,db}$ ($q = 0$)	$O_{i,db}^{\mathbf{A}}$	$\langle 110 \rangle$	2.59	1	3.63
	$O_{i,db}^{\mathbf{B}}$	$\langle 001 \rangle$	3.20	2	1.77
	$O_{i,db}^{\mathbf{C}}$	$\langle 111 \rangle$	2.90	8	0.87
	$O_{i,db}^{\mathbf{D}}$	$\langle 110 \rangle$	3.43	2	0.70
O_i ($q = -2$)	$O_i^{\mathbf{A}}$	$\langle 001 \rangle$	3.20	2	0.62
	$O_i^{\mathbf{B}}$	$\langle 111 \rangle$	4.73	8	0.14
V_{Sn} ($q = -4$)	$V_{Sn}^{\mathbf{A}}$	$\langle 001 \rangle$	3.20	2	4.14
	$V_{Sn}^{\mathbf{B}}$	$\langle 111 \rangle$	3.71	8	3.25
	$V_{Sn}^{\mathbf{C}}$	$\langle 100 \rangle$	4.73	4	2.23
	$V_{Sn}^{\mathbf{D}}$	$\langle 101 \rangle$	5.71	8	2.30
Sn_i ($q = +4$)	$Sn_i^{\mathbf{A}}$	$\langle 001 \rangle$	3.20	2	0.76
	$Sn_i^{\mathbf{B}}$	$\langle 100 \rangle$	4.70	2	0.68
$[Sn_i - V_O]$ ($q = +4$)	diss.	$\langle 001 \rangle$	3.20	1	1.78

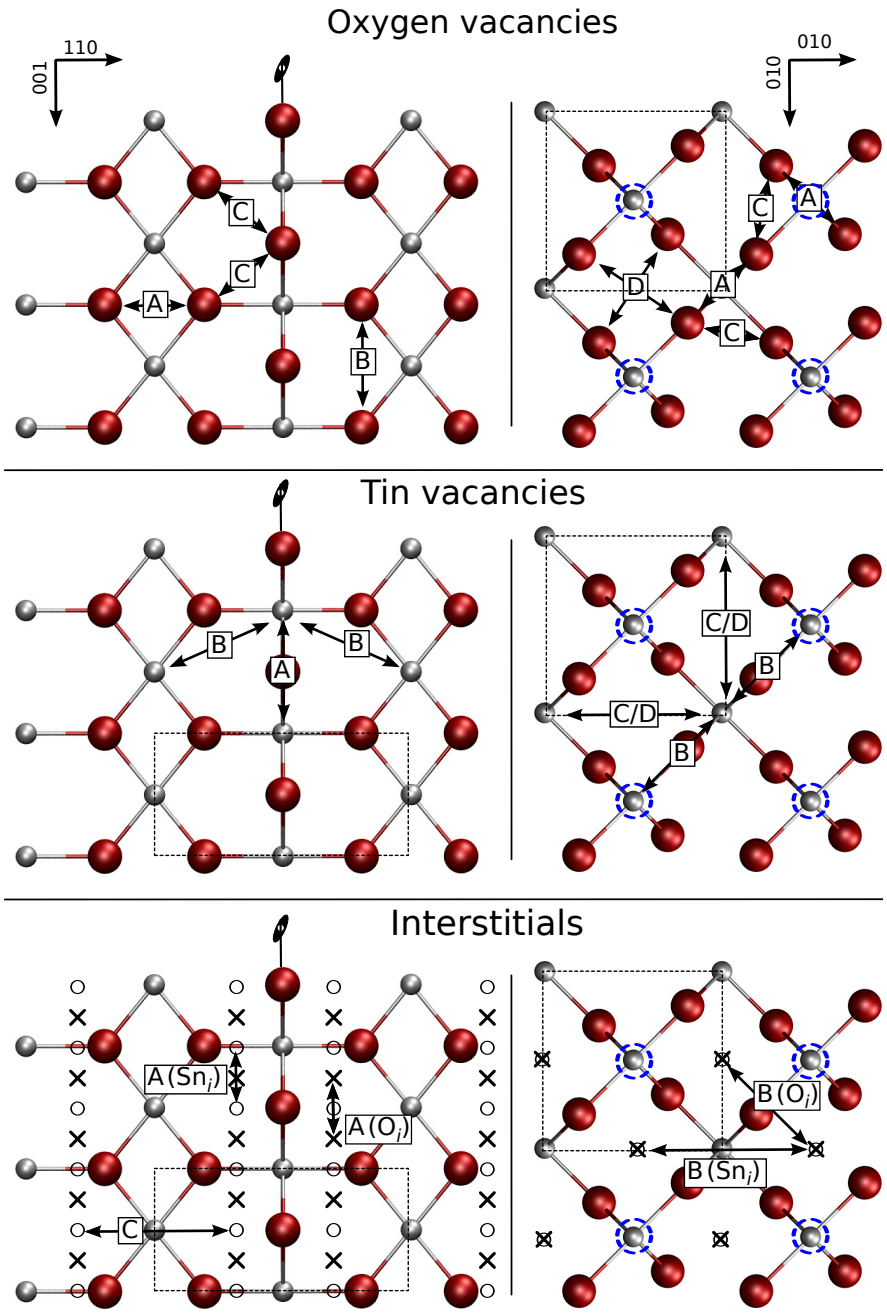


Figure 52: Rutile lattice projected onto {110} and {001} planes. Diffusion mechanisms are labeled and the corresponding energy barriers can be found in Table 12. The circled tin atoms differ in elevation ($c/2$) perpendicular to the image compared with the uncircled ones. Little crosses and black circles mark the locations of the tetrahedral and octahedral interstitial sites, respectively. These sites are located in the front or the back of the atomic layer shown on the left side.

10.2.1 *Vacancies*

Oxygen Vacancy

From the preceding discussion of the defect thermodynamics in SnO_2 (chapter 7) it follows that the oxygen vacancies are the most abundant intrinsic point defect, which can occur in the charge states $q = 0$ and $q = +2$. Thus, migration energies need to be evaluated for these two charge states, separately.

The results listed in Table 12 show that the migration energies of the neutral vacancy are for all jump directions by > 0.3 eV higher than those of the doubly charged defect. This is due to the binding effect of the vacancy state which is occupied for V_O^0 but not for V_O^{2+}. The hierarchy of the energy barriers for the various jump directions, however, is not that much affected by the charge states as it is also depicted in Fig. 53. The trend of increasing migration energies for V_O^0 is also observed for other TCOs like ZnO [289] and In_2O_3(chapter 9).

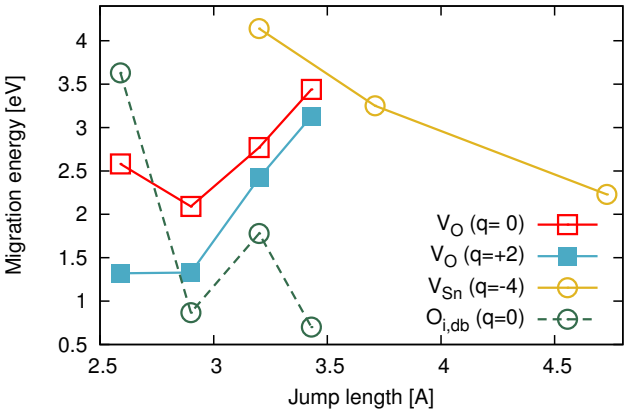


Figure 53: A graphical representation of the barrier data in Table 12.

For SnO_2, the diffusion of oxygen vacancies is determined by atomic jumps into the second neighbor shell ($E^M[V_O^{\mathbf{C}}, q = +2] = 1.33$ eV) as depicted in Fig. 52(top). Vacancy jumps to the nearest neighbor oxygen site (process $V_O^{\mathbf{A}}$) is a motion directly in-between two tin ions. For the neutral vacancy this migration path has a higher energy when compared to the second neighbor jump ($V_O^{\mathbf{C}}$) whereas the two processes have about the same energy in the charged case. The first neighbor process has a multiplicity of one compared to eight for the second neighbor case. Therefore, only the processes via the second neighbors can constitute a pathway throughout the whole crystal. The favorable first neighbor process increases the effective multiplicity from eight to nine in the case of the charged vacancy.

Jumps to lattice sites beyond the second nearest oxygen atoms have significantly higher migration energies (> 2 eV, processes $V_O^{\mathbf{B}}$ and $V_O^{\mathbf{D}}$) even for V_O^{2+}. Since these diffusion paths are structurally not blocked by other ions the increased migration energy for longer distance jumps (Fig. 53) is only due to electrostatic energy differences

between the saddle point and the ground state*. Fig. 53 also illustrates the anomal behavior of the nearest neighbor process as a result of the obstruction by tin.

Oxygen double vacancy

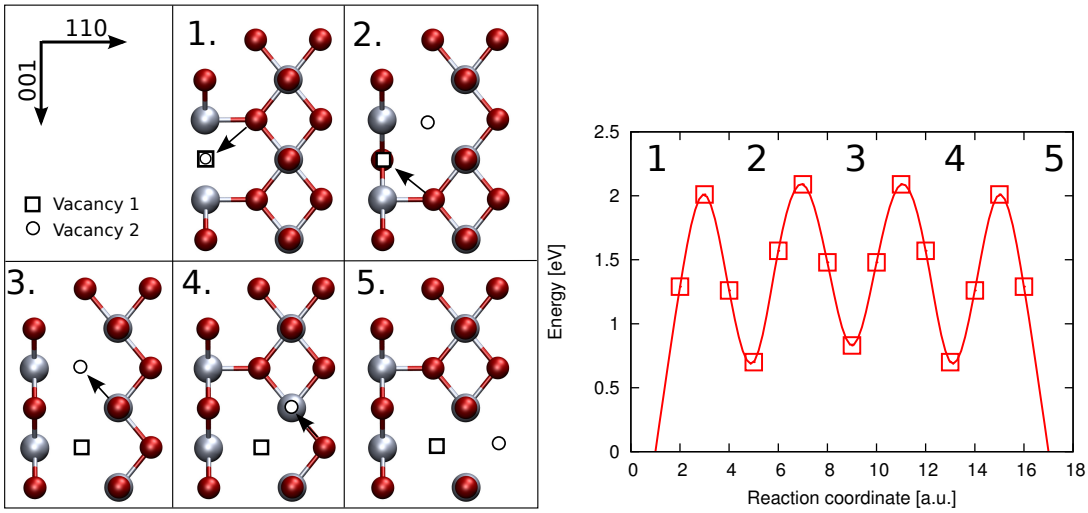


Figure 54: (left) Low energy migration path for an oxygen double vacancy in the charge state q=+2. After five subsequent jumps the double vacancy attains its most stable bound configuration. (right) The energy profile for the double vacancy migration. The local minima correspond to other bound configurations of the double vacancy with lower binding energy.

Following the discussion on point defects in SnO_2 (chapter 7) the double oxygen vacancy can exist in two energetically stable arrangements but with only one predominant charge state. Additionally, there is a series of weakly bound geometries. The first observation is that the two strongly bound configurations cannot transform into each other by simple atomic displacements. Therefore, other more weakly bound geometries are relevant as transition states for the migration of the double vacancy from one stable configuration into another as depicted in Fig. 54. The low energy path of a double oxygen vacancy involves a sequence of four jumps of type V_O^C (lowest energy path for single V_O). We have stepwise calculated the migration energy for one vacancy in the presence of a second vacancy, both forming the initial associate. The thus obtained migration barriers are slightly smaller $E^M \sim 1.2$ eV when compared with $E^M = 1.3$ eV for the single vacancy jumps of the same charge state $q = +2$. The variation for the different steps within the sequence is small (~ 0.1 eV) and also shown

*The linear behavior is expected by reformulating the saddle point configurations as $V_O^{2+} = O_i^{2-} + 2V_O^{2+}$ and $V_O^0 = O_i^{2-} + 2V_O^+$ and considering the electrostatic cohesion of this complexes as a function of distance.

in Fig. 54. For the first event in the sequence, however, the difference in binding energies between the stable and meta stable configurations of the bound double vacancies has to be overcome ($\sim$ 0.7 eV) leading to an overall activation energy of $\sim$ 2 eV. This process is, therefore, less likely than oxygen diffusion through single vacancy jumps of V_O^{2+}, however, energetically comparable with jumps of V_O^0.

Tin Vacancy

According to the discussion of the thermodynamic stability of the various point defects in chapter 8 the equilibrium concentration of V_{Sn} should be very low. Therefore, the presence of V_{Sn} would be an indicator of a non-equilibrium situation due the limited mobility of this defect species. In return, the calculated migration energies can be used to estimate the actual temperature limit up to which a non-equilibrium concentration can be maintained within the material. The migration of cation vacancies is also of major technological importance since cation substituting impurities (antimony) are often used as dopants [101]. The stability and kinetics of the cation dopants is majorly determined by the presence and migration of the cation vacancies. For the cation vacancy we have considered four different migration processes which are depicted in Fig. 52 with the corresponding data in Table 12. Processes $V_{Sn}^{\mathbf{C}}$ and $V_{Sn}^{\mathbf{D}}$ are indicated with the same arrows in the figure and correspond to in-plane and out-of-plane movements, respectively. The intermediate position of the migrating tin is located exactly at the interstitial position and the total jump distances are $\lambda(V_{Sn}^{\mathbf{C}})$=4.73 Å and $\lambda(V_{Sn}^{\mathbf{D}})$= 5.74 Å, respectively. It is remarkable that jumps to the third and fourth neighborhood are those with the lowest migration energy, $E^M[V_{Sn}^{\mathbf{C}}, q = -4] \sim 2.23$ eV. Jumps to the first and second neighborhood ($V_{Sn}^{\mathbf{A}}$, $V_{Sn}^{\mathbf{B}}$) have significantly higher migration energies due to close encounters with oxygen atoms. This anomalous trend is also visualized in Fig. 53 showing the migration barriers E^M as a function of the jump length. The reason for the lower migration energies of the long jumps is that the intermediate position corresponds to a high-symmetry interstitial site, which provides a bulk-like coordination and leading also to an intermediate local minimum for the migrating cation. The depth of this local minimum is $\sim$ 0.6 eV with respect to the saddle point energy. This minimum is sufficiently deep to temporarily capture the cation. Therefore, the effective jump lengths of these processes are only $\lambda(V_{Sn}^{\mathbf{C}})$=2.37 Å and λ ($V_{Sn}^{\mathbf{D}}$)=2.82 Å. The rate controlling step for tin vacancy migration is the displacement of a lattice cation into an interstitial position in close proximity of the vacancy. The considerably deep local minimum can lead to dynamical correlation effects and therefore reduces the effective jump rate.

10.2.2 *Interstitials*

Tin Interstitials

According to the results presented in chapter 7 tin interstitials (Sn_i) can never be majority defects in thermodynamic equilibrium. However, Sn_i is the major defect to mediate tin diffusion and can furthermore influence the conductivity of the material. Tin interstitials also occur as a result of the sputtering process under non-equilibrium conditions.

For Sn_i diffusion we have considered both, regular interstitial motion along the channel-like structures in *c*-axis direction as well as migration within the $a-b$-plane. As we have described in section 7.2.1, there are two different interstitial configurations. The $a-b$ in-plane motion can be described by a subsequent transformation between these two defect configurations (process $Sn_i^{\mathbf{B}}$) and is essentially an intersticialcy mechanism with additional local minima. The motion along $\langle 001 \rangle$ is via the channel structures and surprisingly has a higher energy barrier ($E^M[Sn_i^{\mathbf{A}}, q=-4] \sim 0.76$ eV) when compared with the in-plane motion ($E^M[Sn_i^{\mathbf{B}}, q=-4] \sim 0.68$ eV). As for the relative stability of the two interstitial configurations, the finite size effects are also particularly important for the migration energies of the interstitial defects which are described in section 10.2.3. Finally, we have calculated the energy barrier for dissolving a $[Sn_i - V_O]$-complex via the interstitial channels. The energy barrier for the dissociation

$$[Sn_i - V_O] \rightleftharpoons Sn_i^{4+} + V_O^0 \tag{10.1}$$

is ($E^{\text{diss}} \sim 1.7$ eV) which closely equals the sum of the binding energy of Sn_i^{4+} and V_O^0 as well as the migration energy along $\langle 001 \rangle$. Since no additional energy contributions arise beside binding and migration energies, we have not calculated other dissociation paths.

Oxygen interstitials

For the oxygen interstitials we have investigated the migration in charge states $q=0$ and $q=-2$. In the first case, the oxygen interstitial is bound to a lattice oxygen site forming a peroxo-dimer. Therefore, the same processes as for the oxygen vacancies are possible (see Fig. 52). Notably, the barriers do not follow the energetic trend of vacancy migration barriers (Table 12). The migration to the largest distance where vacancy migration is most unfavorable provides the lowest energy path in the case of the dimers (process $O_{i,db}^{\mathbf{D}}$). This anomalous trend is also depicted in Fig. 53 and results due to the covalent bond between the interstitial and lattice oxygen. For process $O_{i,db}^{\mathbf{D}}$ the oxygen sites are directly facing each other along the dimer axis providing an ideal geometry in order exchange the surplus interstitial between two lattice oxygen sites. For the other cases the dimer has to be rotated resulting in higher migration energies. It is also noteworthy that process $O_{i,db}^{\mathbf{D}}$ is only a motion within the $a-b$ plane. For

isotropic diffusion it is necessary that additionally $O^{C}_{i,db}$ is activated with a slightly higher energy.

In its negative charge state the interstitial is located on the tetrahedral interstitial sites marked with crosses in the Fig. 52. The negative oxygen interstitial is large and causes a significant lattice relaxation around the defect. Especially, lattice oxygen atoms are pushed strongly outwards resulting in a saddle point-like geometry for the intersticialcy mechanism. Consequently, the calculated migration energy for intersticialcy migration is very low ($E^M < 0.15$ eV). In contrast, the jumps in perpendicular direction within the channels along $\langle 001 \rangle$ have a larger migration energy ($E^M = 0.62$ eV).

10.2.3 *Finite size effects*

From the calculated migration energies we can now extract the rate controlling processes, which are given in Table. 13 for diffusion parallel and perpendicular to the c-axis of the tetragonal crystal structure. For vacancies there is no distinction, because the same process leads to a an almost isotropic diffusional motion. This analysis is valid under the assumption that finite size effects only cause an offset in calculated migration energies, but do not change their hierarchy. Test calculations, however, revealed that this assumption is not necessarily justified when interstitial and intersticialcy processes are compared. Thus, we have treated these cases, separately. The scaling behavior with supercell size is shown in the Fig. 55 for the processes listed in Table 13 and three different cell sizes (72, 162, 216). We have also assumed that the regular scaling behavior should be as $1/N$, typical for elastic interactions and treat exceptions as anomal.

Table 13: The rate controlling processes for each defect are listed in lines 1 and 2 for the directions perpendicular and parallel to the c-directions and with the corresponding barriers (eV) in lines 3 and 4, respectively.

Defect	V^{0}_{O}	V^{2+}_{O}	$(V_O - V_O)^{2+}$	V^{4-}_{Sn}	O^{2-}_{i}	$O_{i,db}$	Sn^{4+}_{i}
process ($\perp c$)	$V^{\mathbf{C}}_{O}$	$V^{\mathbf{C}}_{O}$	$V^{\mathbf{C}}_{O}$	$V^{\mathbf{C}}_{Sn}$	$O^{\mathbf{B}}_{i}$	$O^{\mathbf{C}}_{i,db}$	$Sn^{\mathbf{A}}_{i}$
process ($\parallel c$)	$V^{\mathbf{C}}_{O}$	$V^{\mathbf{C}}_{O}$	$V^{\mathbf{C}}_{O}$	$V^{\mathbf{D}}_{Sn}$	$O^{\mathbf{B}}_{i}$	$O^{\mathbf{D}}_{i,db}$	$Sn^{\mathbf{B}}_{i}$
scaled ΔE^M ($\perp c$)	1.7(7)	1.2(8)	2.0(0)	2.2(7)	0.1(4)	0.6(9)	0.0(8)
scaled ΔE^M ($\parallel c$)	1.7(7)	1.2(8)	2.0(0)	2.1(2)	0.1(4)	0.8(8)	0.8(3)

For V^{2+}_{O} and $O_{i,db}$ the migration energies are not affected by the cell size. The correction due to the cell size is smaller than the estimated accuracy with respect to the exchange-correlation functional. In contrast, jumps of V^{0}_{O} show a pronounced size effect which is related to the so-called band-filling effect [192]. Due to the artificially high concentration of defects in a supercell geometry the defect induced states disperse

depending on cell size. The position of the defect state is, however, also a function of the reaction coordinate along the diffusion path and leads to a different dispersion in the ground-state and the transition-state. This effect is especially large when the defect state crosses the host bands of the material at the saddle point which is the case in SnO_2. As a result the barriers for V_O^0 are overestimated. Further, because the defect level position is relevant for diffusion of this defect we have also assumed a stronger effect due to the chosen exchange-correlation functional, or more generally of the band-gap. In order to test these ideas, we have recalculated the barrier for the neutral charge state with the PBE0 functional and found an increase of the migration energy by 0.8 eV relative to the unscaled LDA value (process V_O^C in Table 12). This can be explained by the fact that for migration the binding vacancy state needs to be broken. Since the defect state of the vacancy is deeper with respect to the CBM when compared with LDA calculations this leads to an increased migration barrier. Based on previous works on sulfides and selenides this effect is not expected to appear for defects without defect induced gap states and was, therefore, not further investigated for other defects [290].

In the case of Sn_i and V_{Sn} the finite size effects are even more pronounced when compared with V_O^0 and are especially large for the two smallest cell sizes. We have found that this is due to the very large jump distances which are comparable to half the supercell width. On the saddle point the defects extend over two lattice sites corresponding to line-defect-like structures with artificial interactions. For the two larger cells this is not the case and we assume a well behaved convergence starting from a cell size of N=162. Therefore, for extrapolating our data we have only used the last two points. In the case of the V_{Sn} we have verified this procedure by calculating another cell size. In the case of the interstitial, we rely on the extrapolation with two data points since the migration barrier is so low that thermal activation would always occur at room temperature even if corrections were applied. For the same reason we have not attempted a scaling for the oxygen interstitial in the negative charge state.

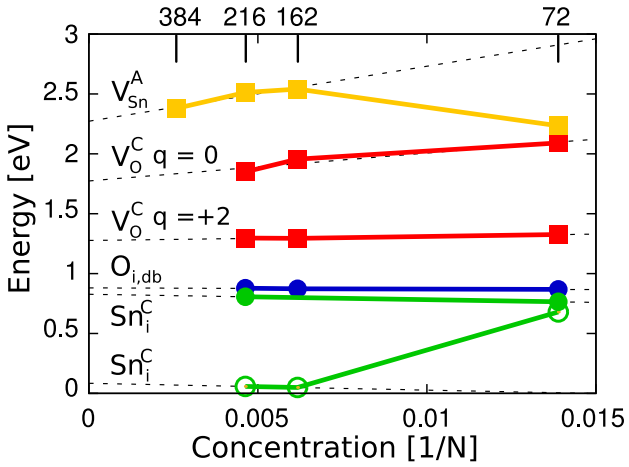


Figure 55: Scaling behavior of migration energies as a function of cell size.

10.3 DISCUSSION AND COMPARISON WITH EXPERIMENT

The finite size corrections have been applied where necessary, and the complete set of diffusion barriers can now be used to interpret experimental data. Due to the sensor activity of SnO_2, however, the comparison can be difficult since surface electron depletion/accumulation effects cannot be distinguished easily from changes of the bulk conductivity. Additionally, it was found for SnO_2 that under more oxygen rich conditions the surfaces have a blocking effect especially for oxygen making it impossible to measure the bulk diffusion under lower temperature [291]. To our best knowledge there is also no diffusion measurement using radioactively marked tracer atoms. Therefore, we have mainly focussed on the experiments of Kamp *et al.* [285, 255] for comparing our results. They measured the chemical diffusion via the electron concentration as a function of the oxygen partial pressure in single crystals and at high temperatures. The latter is important to avoid the blocking by the surfaces. Kamp *et al.* report an average activation energy of 1.1 eV $\pm$ 0.3 eV for the diffusing (positively charged) species. They assume that V_O^{2+} is the mobile defect species, which is confirmed by our calculated value of 1.2(8) eV $\pm$ 0.1 eV. The absent anisotropy in both, theoretical and experimental data supports this interpretation. The active migration mechanism for oxygen vacancies is along $\langle 111 \rangle$ with components of the jump vector in all crystallographic directions. The resulting anisotropy is, therefore, very small and only affects the prefactor due to the c/a-ratio.

This interpretation should be reliable with respect to other explanations. Alternative mobile species could be the tin interstitial Sn_i^{4+} as well as the double vacancy $(V_O - V_O)^{2+}$. Their migration energies are, however, 0.19 eV $\pm$ 0.1 eV and 2.01 eV $\pm$ 0.1 eV and are obviously not in line with the experimental data, even if the uncertainties of the calculated values are taken into account. Other possibilities for interpreting the experimental result include the migration of hydrogen defects such as H_O ($E^M \sim$ 2.6 eV) [132] as well as the dissociation reaction of H_O ($E^{\text{diss}} = 1.4$ eV [253]. While the first value is too large, the second process would not lead to an oxygen partial pressure dependence of the conductivity which was found in the diffusion studies.

The dissociation of Eq. 10.1 ($E^{\text{diss}} \sim 1.7$ eV) has also a rather high activation energy. In addition, the formation energy of the $[Sn_i - V_O]$-complex is too high (chapter 7) and not even expected to be bound in especially acceptor doped material since both, the isolated V_O and Sn_i are both charged in that case. The samples of Kamp *et al.*contained intentional acceptor impurities.

The migration of the oxygen interstitials can be excluded at high temperatures because of their high formation energy. The experiments have been conducted at rather reducing conditions so that the concentration of oxygen dumbbells should be negligibly small. Finally, the V_{Sn} have too high formation and migration energies to be responsible for any of the experimental observations.

Interestingly, Kamp *et al.* and also Goepel *et al.* have observed some anomalies of the defect chemistry which they attribute to the presence of tin vacancies [285, 255, 115]. The "missing" defect appears to have a slower kinetics and a stronger partial pressure

dependence. While we can only speculate about the nature of this defect the best candidate according to our results is the $(V_O - V_O)^{2+}$ but not the tin vacancy. We have shown in section 7.2.1 that different agglomerates of donor defects are possible. Also, in the literature larger planar oxygen deficient structures have been observed [292] and modeled using electronic structure methods [72]. These defects can be thought of as larger clusters of oxygen vacancies and similar to the Magneli phases known for Rutile [293]. The formation and decomposition of these more extended defect structures are a plausible source for the reported anomalies. Due to the dual valence of tin, sub oxides do exist and thus, there is a natural tendency for these agglomeration reactions.

The lowest diffusion constants are obtained for oxygen interstitials in the negative charge state. This charge state is, however, only stable at extremely high Fermi energies and the concentration is always very low. Slightly higher concentrations can be expected for the oxygen dumbbell with a somewhat larger migration energy. However, both defects will already anneal at slightly elevated temperatures and non-equilibrium concentrations are not expected.

In fact it was found, that sputtering of antimony doped tin-oxide (ATO) in oxygen atmosphere does not lead to a decrease of conductivities in SnO_2 [138]. During this process, oxygen interstitials are implanted within the subsurface which could lead to a non-equilibrium concentration of oxygen interstitials. In the case of SnO_2 our results indicate that the oxygen interstitials anneal already during the deposition process in agreement with experimental observations. A rather different behavior is obtained in the case of In_2O_3, where both, the formation energy of interstitials is lower and the migration energies are higher [294]. As a result non-equilibrium interstitial concentrations can be achieved in In_2O_3, but not in SnO_2.

For interstitials and in contrast to the vacancies we generally find a slight anisotropy for diffusion with the easy direction perpendicular to the *c*-axis. This is somewhat counterintuitive since the channel like structures extend along *c*-direction. The elastic strain, however, causes the ground state of the interstitial defects to assume a conformation which is especially suitable for intersticialcy migration perpendicular to the *c*-axis. In result, the diffusion is faster in these directions for both, oxygen and tin interstitials.

Finally, for V_{Sn}^{4-} we find higher migration energies when compared with charged oxygen vacancies. Considering the relative abundance of the two defect types, the diffusion of tin vacancies should be negligibly small under any condition compared with oxygen vacancy diffusion. This has also implications for the doping of the material. Cation dopants should be kinetically more stable at non-equilibrium concentration when compared with anion substituting dopants. By the same token V_{Sn} could be present at non-equilibrium concentrations up to relatively high temperatures.

10.4 SUMMARY AND CONCLUSION

We have investigated the migration processes of intrinsic point defects in SnO_2. For vacancies we have found that nearest neighbor processes are not relevant. This is especially significant for the tin vacancies where migration does only occur to third and fourth neighbor sites via an intermediate minimum at the interstitial sites. There is a strong charge dependent migration energy for the oxygen vacancies with a faster migration in the charged case. The oxygen interstitials in negative charge states have negligible migration energies and will diffuse out of samples even at low temperatures and are not an important defect for SnO_2. The oxygen dumbbell defects do not influence the conductivity of the material and anneal already at slightly elevated temperatures.

Tin interstitials have only little lower formation energies and should according to the present study also be a fast diffusing species. In addition, they may be trapped at low temperatures at oxygen vacancies but only in a non-equilibrium situation. Otherwise such defect complexes have too high formation energies and dissociation should be possible at higher temperatures.

The migration energy values for the charged oxygen vacancies are in best agreement with experimental single crystal data within the error bounds. We have obtained a negligible anisotropy for diffusion of oxygen vacancies which agrees well with the measurements. Additionally, we have pointed out the relevance of double oxygen vacancies as plausible source for anomalies in the defect equilibria of the material.

DEFECT KINETICS: In_2O_3 VS. SnO_2

We are now in the position to compare the defect kinetics of the two materials SnO_2 and In_2O_3. In the following we summarize the key findings and list important similarities and differences between the two materials.

The major similarities are:

- Oxygen diffusion is mediated via oxygen vacancies at reducing conditions.
- V_O^0 migrates more slowly when compared with V_O^{2+}.
- Cation vacancies migrate generally more slowly than oxygen vacancies. This trend contrasts, however, with the behavior of other TCOs like ZnO [289, 295, 239] and CdO or Ga_2O_3 [108].
- For both materials intersticialcy migration is dominant over regular interstitials migration.
- Vacancy migration is isotropic for both materials[†].

More interesting are, however, the differences between the two materials:

- Charged oxygen vacancy migration is slightly faster in In_2O_3.
- Cation vacancy migration is faster in In_2O_3[‡].
- Cation and anion interstitials are fast diffusers in SnO_2 but slow in In_2O_3. In the case of In_2O_3 interstitials migrate even slower when compared with V_O^{2+}.
- Oxygen dumbbell diffusion is slightly faster in SnO_2.
- Lowest energy barriers determine the effective migration energy in SnO_2 but not in the case of In_2O_3.
- Interstitial migration in SnO_2 is anisotropic especially for Sn_i^{4+}.
- Light induced enhancement of diffusion is expected for In_2O_3 but not in the case of SnO_2 due to low abundance of O_i^{2-}.

[†]Tin oxide is tetragonal.

[‡]This is true when finite size effects are considered for In_2O_3.

Part V

SURFACES

The following part of the thesis finally deals with the surface of SnO_2 and In_2O_3. Effects like intrinsic doping and stoichiometry variations, etc. occur on surfaces in a similar way as in the bulk. At surfaces, however, the corresponding equilibria and chemistry can substantially differ.

The first chapter of this part is devoted to surface properties of SnO_2. SnO_2 is a good model system, because a large body of data exists in literature. Despite extensive research, the nature of SnO_2-(110) reconstructions remained unknown. Therefore, we focus mainly on this surface and complement preexisting data. Our main result is a revised surface phase diagram containing surface terminations consistent with all published experiments.

The second chapter deals with the surfaces of In_2O_3. Since only little is known on a microscopic level, our investigations of In_2O_3 survey a number of surfaces. We investigate stoichiometry variations and effects of water adsorption, surface stress and doping. We propose surface phase diagrams for (111), (011), (001) and (211) surfaces and discuss how surface terminations of (001) faces change at high n-type doping levels such as in ITO. The presented results are compared with existing STM measurements.

We conclude by pointing out important similarities and differences between SnO_2 and In_2O_3.

11

SURFACES OF TIN OXIDE

11.1 INTRODUCTION

The surfaces of SnO_2 have been thoroughly investigated in the past by low energy electron diffraction (LEED) [296, 297, 298, 299, 300, 301, 302, 303, 304], ion scattering spectroscopy [302, 298, 305], Auger electron spectroscopy, scanning tunneling microscopy (STM) as well photo electron spectroscopic techniques (UPS,XPS)[51, 302, 138, 306] and non-contact AFM (nc-AFM)[301]. The interpretation of the experimental data was supported by numerous first-principles studies. [307, 308, 309, 310, 311, 312, 298, 313] The energetic order of the surface stability has been established as $(110) < (100) < (101) < (001)$ [309] for bulk truncated surfaces. Moreover, all low index surfaces have been investigated under oxidizing as well as reducing conditions. It was found that all surfaces can be reduced, if the oxygen chemical potential is appropriately chosen. Generally, on stoichiometric surfaces, (1×1) reconstructions are observed corresponding to simple bulk truncations.

Under reducing conditions all surfaces undergo a transition where (100), (101) and (001) surfaces remain with a (1×1) reconstruction. Occasional observations of a (1×2) reconstruction of (101) surfaces could be rationalized with corresponding DFT calculations [298].

While the latter surfaces exhibit fairly simple reconstructions in the stoichiometric as well as in the reduced state, the (110) surface shows an unexpected richness of different structures and surface defects, some apparently being even metastable but nevertheless highly reproducible [298]. Besides the bulk truncated stoichiometric surface a series of reconstructions have been observed in LEED, but also in atomically resolved STM [298, 299, 300, 301, 302, 303, 304]. Depending on oxygen pressure, temperature but also the preparation condition and sample history $c(2 \times 2)$, (4×1), (1×1) (different from stoichiometric), (2×1), and (1×2) reconstructions were identified and copiously reviewed in Ref. [51]. Despite this large variety of surface reconstructions, only two appear to be true ground-state structures as a function of oxygen chemical potential. Although high temperature (1×1)-phases have been proposed by first-principles modeling [314, 312, 310], they conflict with experimental observations.

It can be concluded that up to now no satisfactory models have emerged for the reduced equilibrium SnO_2-(110) surfaces. Taking this as a starting point, we have conducted a survey of oxygen deficient surface reconstructions by calculating the surface tensions of a series of over 200 structures using total energy calculations based on density functional theory (DFT). We show the existence of a wealth of different surface reconstructions which to our believe should also be observable in experiments.

The calculations are based on the local density approximation (LDA). Details about the computational setup and the convergence properties can be found in Ref. [A9]. The slabs always contained at least 12 atomic layers or 4 trilayers of SnO_2 with approximately 192 atoms. The lateral extent of the cell was a (4×2) super-mesh in order to accommodate all experimentally observed structures. The low index surfaces of SnO_2 have no net dipole moment so that it was always possible to use symmetric slabs.

For STM image simulations we have used the standard approach proposed by Tersoff and Haman [315]. For these calculations the *k*-point density was increased in order to obtain a representative Brillouin zone average. Since in experiments exclusively the empty states were mapped, we have integrated the states towards positive energies with respect to the Fermi energy of the system. The image height was adjusted in order to obtain contrast of reasonable detail. The surface tensions as a function of the oxygen chemical potential diagrams were obtained via Eq. 4.40 of the formalism described in section 4.6.

11.2 SnO_2 SURFACES

Tin can exist in the oxidation sates Sn^{4+} and Sn^{2+}. In the case of Sn^{2+} the Sn-$5s^2$ electrons are localized and form a lone-pair. This provides a possibility for the formation of sub-oxides (*e.g.* litharge-SnO). In the case of surface reconstructions the formation of local sub-oxides is a means to establish stable and charge compensated surfaces.

Except for SnO_2-(001), all low index SnO_2 surfaces are type-II [(001) is of the first kind] following the classification scheme by Tasker [213, 214]. This means that the stacking sequence perpendicular to the surfaces consists of symmetric trilayers without any net dipole moment. Therefore, the bulk truncated SnO_2-surfaces are stable configurations which can exist without any reconstruction or stoichiometry variation. Experimentally both, stoichiometric and oxygen deficient surfaces are observable reflecting the oxygen deficient nature of this material.

Depending on the oxygen chemical potential (oxygen partial pressure and temperature) (110), (100), (101) and (001) surfaces exhibit two regimes. This behavior is illustrated in Fig.56, where the surface tension is plotted as a function of the oxygen chemical potential (μ_O) for these surfaces. For each surface a transition from stoichiometric to reduced surfaces occurs. The reduction point of the surfaces scales with their

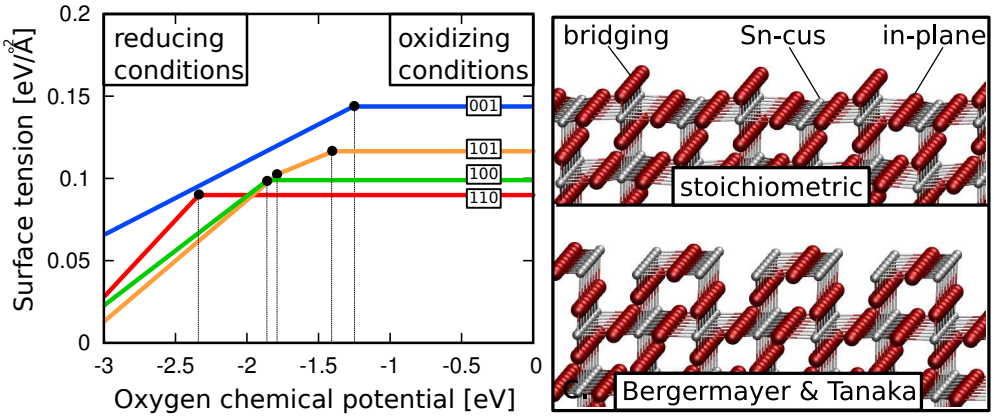


Figure 56: (top) Surface phase diagrams for (110), (100), (101) and (001) surfaces as calculated within the LDA. (bottom) Geometries of the stoichiometric and a highly reduced (110) surfaces.

absolute stability. More stable stoichiometric surfaces have a reduction point at lower chemical potential.

Generally, chemically sensible reduction of surfaces is possible if there is at least one atomic species, which is effectively reduced. In SnO_2, Sn undergoes the transition from Sn^{4+} to Sn^{2+}. It is important to note that the localization of the lone-pair is stereo-active *i.e.* its localization will induce a break of symmetry. In fact, the calculations show that the lone-pair only localizes in two distinct geometries. Either it localizes in a four-fold square-pyramidal coordination as it is found in the litharge structure of tin-monoxide (SnO) or in a three-fold trigonal-pyramidal arrangement corresponding to SnO-massicot or Pb_3O_4. By inspecting the different SnO_2 surfaces it is apparent that removing topmost oxygen from (100), (101) and (001) surfaces results in the trigonal-pyramidal configuration of Sn. From Bader analysis the charge localization is confirmed and thus obtained surfaces have low surface tensions and appear as reduced surfaces in the phase diagram (Fig. 56 at lower μ_O.

In this context, the (110) surface is an exception because removing the topmost oxygen (bridging-O) results in 4-fold coordination with no pyramidal arrangement. Bader analysis confirms that no lone-pair localization occurs in this case. Only if every second in-plane oxygen (second O-layer) is removed, a lone-pair localization within a trigonal-pyramidal arrangement can occur. The resulting structure is shown at the bottom of Fig.56 and has been previously proposed by Bergermayer and Tanaka [310].

A comparison of the thermodynamics of surfaces from DFT calculations with experimental findings shows good agreement for (100), (101) and (001) surfaces, whereas, several discrepancies can be observed for the (110) orientation which shall be summarized in the following.

Up to a a chemical potential of ~ -1.3 eV all surfaces are stoichiometric corresponding to a temperature of $T \sim 500$K under ultra high vacuum conditions (UHV).

This finding is in excellent agreement with measurements of the surface composition [302, 302, 305]. Further, the loss of oxygen on (100) and (101) is found at comparable oxygen chemical potential also in agreement with experimental results. On the other side reduction of the (110) surface is found at slightly lower temperatures when compared with (101) and (100) which is at odds with the calculated phase diagrams.

A highly oxygen deficient ordered (1×1) reconstruction with *pm* symmetry was proposed using DFT [310]. Experimentally it was found, however, that the highly reduced (1×1) reconstruction exhibits a high degree of disorder. In addition, the high temperature (1×1) reconstruction was found to have a local (2×1) periodicity [304] and step heights corresponding to fractional interlayer distances [305] conflicting with the proposed geometry. Also, the calculated transition temperature is too high when compared with experiments.

Besides the (1×1) reconstruction an (1×2) reconstruction exists at even higher temperatures with a large corrugation and a significant disorder for which no thermodynamically stable model could be found by means of electronic structure calculations.

In addition, $c(2 \times 2)$ and (4×1) reconstructions are frequently observed [297, 299, 300, 51, 298, 301, 302, 303, 304]. Two findings hint to the metastable character of these structures. Firstly, the oxygen content increases upon heating of these surfaces. Secondly, in reversibility studies it was shown that these reconstructions cannot be reformed if no sputtering is employed. The $c(2 \times 2)$ reconstruction appears to be formed simultaneously or prior to the formation of the (4×1) reconstruction. A coexistence of both was frequently observed pointing also towards non- or partial equilibrium of the surface in these cases [51]. For the (4×1) reconstruction we have found three different models in the literature [303, 299, 305] whereas only one contribution has discussed the nature of the $c(2 \times 2)$ surface [303].

In summary, good agreement of theory and experiment is found for all SnO_2 surfaces except for the (110) oriented one. For all (110)-reconstructions observed in experiments other than the stoichiometric surface no satisfactory models exist (reduced-(1×1), (1×2)). The same holds true for the metastable $c(2 \times 2)$ and (4×1) reconstructions .

11.3 RESULTS AND DISCUSSION

11.3.1 *Weakly reduced (110) surfaces*

First, we discuss the effect of mild reduction of the stoichiometric (110) surface. In this case, a reduction on a microscopic scale is expected by the formation of surface point defects. The clustering of those defects may subsequently lead to new surface reconstructions. In turn, the predominant surface defects hint to the nature of the reconstructions which occur at more reducing conditions.

In Fig. 57 the formation energies of different point defects within the surface are shown as a function of the oxygen chemical potential where the stoichiometric surface

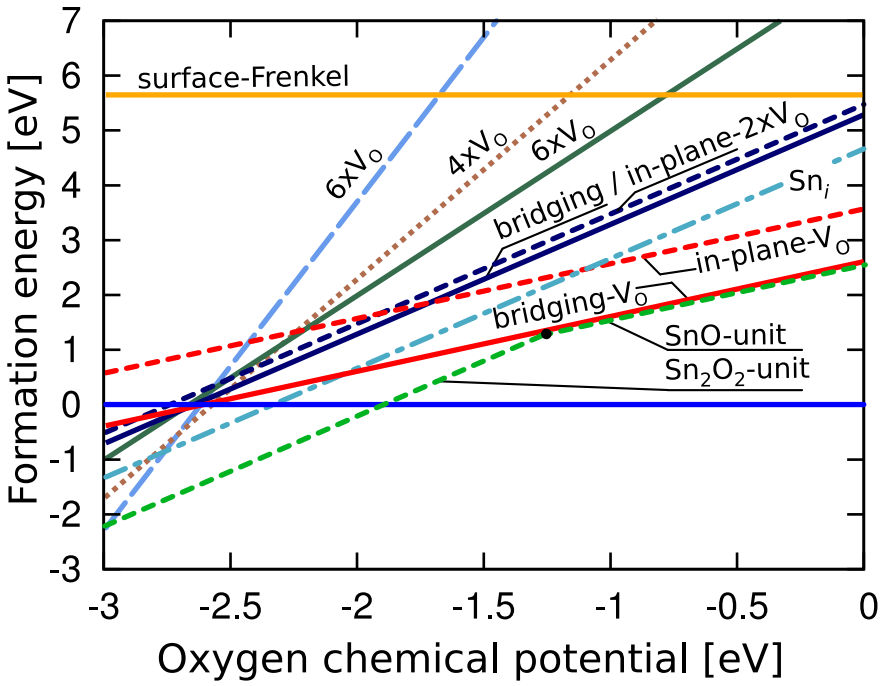


Figure 57: Formation energies of surface defects in the stoichiometric SnO_2-110 surface.

is the reference. For simplicity, the calculated formation energies are obtained for the charge neutral case. Because electronic defects are not explicitly considered, the presented numbers represent upper bounds for the formation energies.

The simplest defect producing oxygen deficiency is the oxygen vacancy in either bridging or in-plane position (see Fig. 56). The formation energy of the first is by ~ 1 eV smaller when compared with the latter. In fact, vacancies in in-plane configuration are dynamically unstable. The barrier for transforming into a bridging vacancy is very small < 0.2 eV impeding their experimental detection.

Other important surface defects are double vacancies, because it is plausible to assume vacancy formation in pairs by ejection of oxygen molecules (single oxygen atoms as surface species are unlikely [316]). The formation energies of double vacancies in the first and second layer are almost identical. This is because of the strong association tendency for in-plane vacancies [308], whereas bridging vacancies show a rather repulsive behavior. In the second layer, several double vacancy geometries exist and the formation energy is lowest for the pair in nearest neighborhood similar as it is in the bulk of the material (see chapter 7).

In contrast to single vacancies, the double vacancy is a dynamically stable species in the second O-layer whereas in bridging position they tend to dissociate. Once an in-plane double vacancy is formed its annihilation is unlikely, since both vacancies

need to be removed at once in order to avoid the existence of a single in-plane vacancy, which is energetically unfavorable.

Alternatively, it is possible to form surface cation interstitials. The formation energy of this defects is, however, high when compared with single vacancies, but lower than for double vacancies. As a matter of fact, surplus interstitials are not available from the gas phase (no gaseous Sn species) nor from the bulk of the material (chapter 7) so that the formation of interstitials can only occur via the formation of surface Frenkel pairs, which have a very high formation energy.

By inspection of the electronic density of states, we find that oxygen vacancies in the first plane are shallow double donors in contrast to vacancies in the second plane. In our calculations, this can be inferred from the fact that electrons occupy host-like conduction band states upon oxygen removal in the first-plane. Also tin surface interstitials are shallow double donors on the surface by assuming the oxidation state Sn^{2+}, whereas Sn interstitial defects in the bulk have a charge state of $q = +4$ presented earlier (chapter 7).

In Fig. 57 the formation energies of several larger vacancy clusters are also shown (only the lowest energy configurations) with, however, relatively high formation energies when compared with the isolated vacancies.

The low formation energies of *e.g.* isolated $V^{2+}_{O,bridging}$ and other positively charged defects implies that stoichiometric SnO_2-(110) surfaces should exhibit an electron accumulation layer at the surface under mild reduction conditions. This is not necessarily observed on other surfaces of SnO_2. The charging of the surface also induces a driving force for charge compensation. It has been observed that stoichiometric surfaces are often rather rough despite of their very low surface tension. This roughening could be a means to compensate the surface charge. If this is the case, smooth stoichiometric (110) surfaces should be obtainable for n-doped SnO_2. Under n-doping the formation of intrinsic donors is less favored and screening is also enhanced.

Within this context we have identified another important surface defect causing oxygen deficiency. This can be thought of as an associated Anti-Schottky defect or SnO-unit. In this case Sn is located on the interstitial site whereas O is positioned above the coordinatively unsaturated Sn site (Sn_{cus}). This defect has a slightly higher formation energy when compared with the bridging oxygen vacancy (the formation energy of the vacancy is overestimated due to long-ranged band-filling super-cell size-effect [192]). More importantly, at lower oxygen chemical potential ($\mu_O < -1.3$ eV) the association of these defects into a symmetrical arrangement as depicted in Fig. 58 (left) becomes favorable. At even lower oxygen potential this Sn_2O_2 defect complex is the lowest energy surface defect on stoichiometric (110)-surfaces.

We stress that in comparison with the formation of vacancies the formation of this defects could be kinetically hindered *e.g.* by the formation of surface Frenkel pairs or cation interstitials as well as surface oxygen, which have altogether high formation energies. The evolution of this defect may occur via reconstruction of vacancy clusters, surface sputtering, or by a splitting from *e.g.* kinks. Since this defect

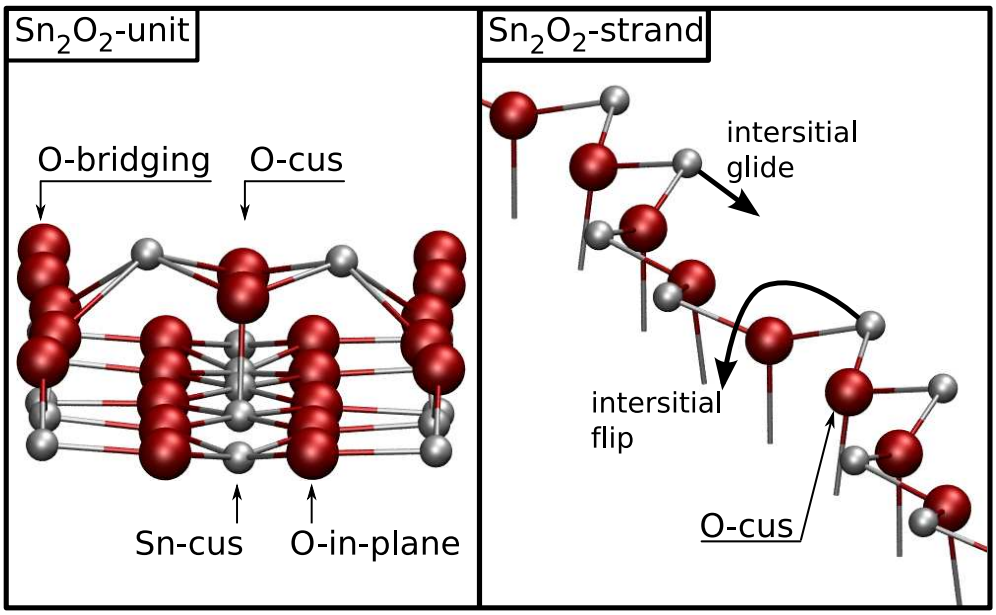


Figure 58: Geometry of the Sn_2O_2-defect (left). The surplus cations bond with two O_{cus} and four bridging oxygen. The more extended version of this defect (right) allows several modes of structural low energy excitations denoted as interstitial flip and interstitial glide.

is charge neutral by forming states at the top of the valence band, we conclude that its formation provides an alternative route for surface reduction on small scales which avoids surface charging in contrast to the formation of surface vacancies.

11.3.2 *Strongly reduced (110)-surfaces*

Disordered (1×1) *reconstruction*

We have taken the novel surface defect described before as a building block for the construction of other possible surface reconstructions. This is achieved by extending the Sn_2O_2 units (Fig. 58) in order to form chains along the $[001]$ direction. This way every second surface interstitial site remains unoccupied and leads to the (2×2) surface structure shown in Fig. 59(a).

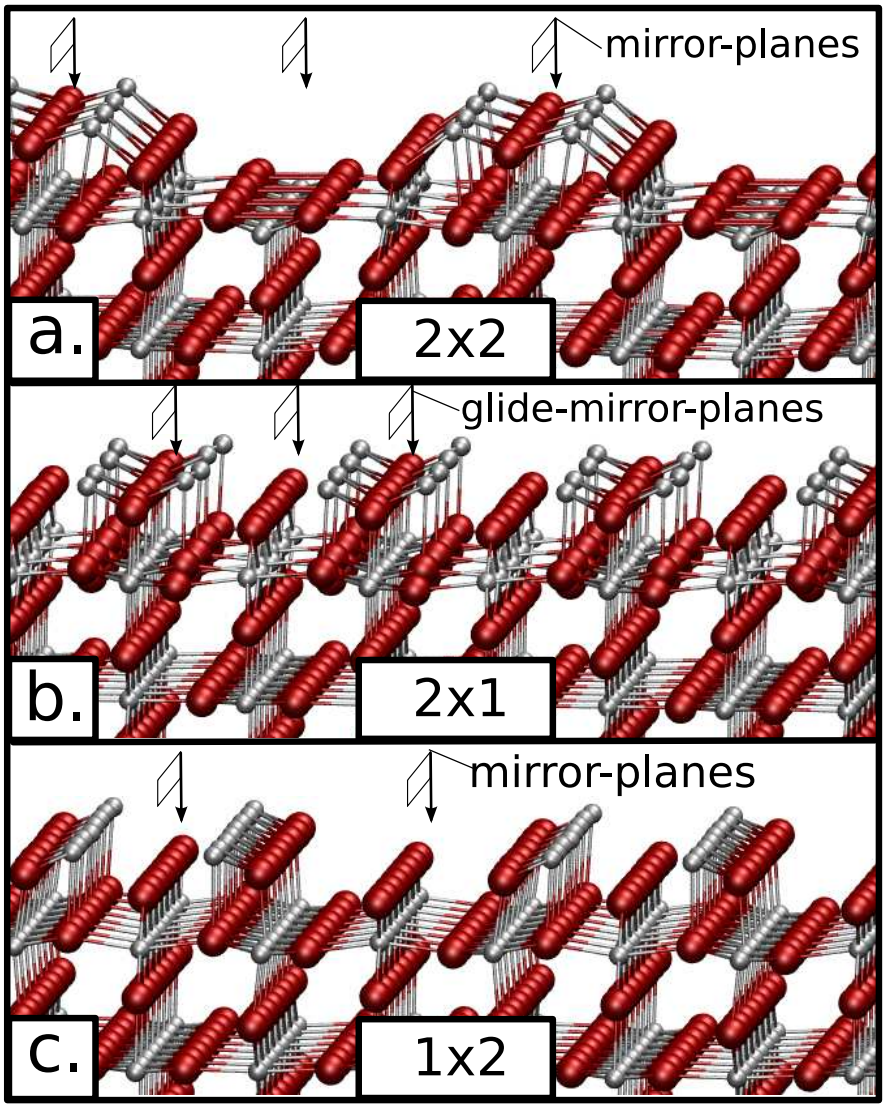


Figure 59: Geometries of reconstructions based on the Sn_2O_2 surface defect. Geometry of (a.) is obtained by multiplying the Sn_2O_2 defect along [001] direction. The geometry in (b.) is the lowest energy structure whereas (c.) is the highest energy configuration.

It is not favorable to form these Sn_2O_2-strands on top of each row of Sn_{cus} because these strands interact with each other along $[1\bar{1}0]$ direction. As it is accentuated in Fig. 58 (left) the Sn_2O_2-defect units form bonds with the neighboring bridging oxygen leading to a geometry similar like in the litharge structure. This attraction to bridging oxygen to the Sn_2O_2-strands conflicts with neighboring strands.

The strands can further be modified by shifting one half of the strand with respect to the other by one lattice vector along $[001]$. This procedure introduces glide mirror planes and leads to the (2×1) zig-zag structure shown in Fig. 59(b). Note, that in this case it is favorable to place strands on each row of Sn_{cus}. It is also noteworthy that in this geometry the tin interstitials do not bond to the bridging oxygen anymore, but to the in-plane oxygen directly below leading to a massicot-type arrangement.

For a full surface coverage with Sn_2O_2-defects, this geometry is also the overall ground-state structure. A striking feature of this reconstruction is, however, that the exact position of the interstitials is not very important on a microscopic scale. Individual interstitials can glide along $[001]$ direction or perform flips to the opposite side of the O_{cus} in $[1\bar{1}0]$ direction at low energy cost of ~ 0.07 eV leading to surface pattern, with other periodicities. These processes are depicted in Fig. 58 (right) starting from a (4×1) pattern which is energetically also very favorable. With these small excitations infinitely many low energy structures can be obtained. The structure with highest energy which can be obtained this way is shown in Fig. 57(c) and has a (1×2) periodicity, where all interstitials are aligned on one side of the O_{cus}.

In Fig. 60 the energetics of this class of surfaces is shown as a function of the oxygen chemical potential. For comparison, the stoichiometric as well as reduced surface of Ref. [310] is also shown. The new surface reconstruction involving tin surface interstitials spans a considerable stability range ($\Delta\mu_O > 0.6$ eV). Besides isolated Sn_2O_2 units at higher oxygen potentials, the (2×2) reconstruction has a narrow stability range. The most stable reconstruction has the (2×1) periodicity as described before and found at slightly lower oxygen potential.

These reconstructions form the ground-state structures but due to the low energetic cost of cation interstitial flip it is not likely that they could be observed experimentally in this pure form. At temperatures where these structures should be observable the interstitial flip is thermally activated and leads to a highly disordered structure which is in continuous transformation, however, maintaining its basic patterns. The high density of states is indicated by several thin black lines in the phase-diagram (Fig. 60) and representing different surface cation arrangements with similar energetics.

A disordered (1×1) reconstruction is observed at about $T = 900\text{K}-1000\text{K}$ experimentally with defect structures which are not well understood. More strikingly it was shown by Sinner-Hettenbach *et al.* [304] that the disordered (1×1) reconstruction exhibits some local order with a (2×1) periodicity and zigzag-like features. The (2×1) zigzag arrangement is also identified as the ground state configuration in the present study. This gives us confidence that we have indeed found the observed structure. Moreover, for the high temperature (1×1) reconstruction fractional interlayer spacings have been observed, which naturally arise for the calculated structures due

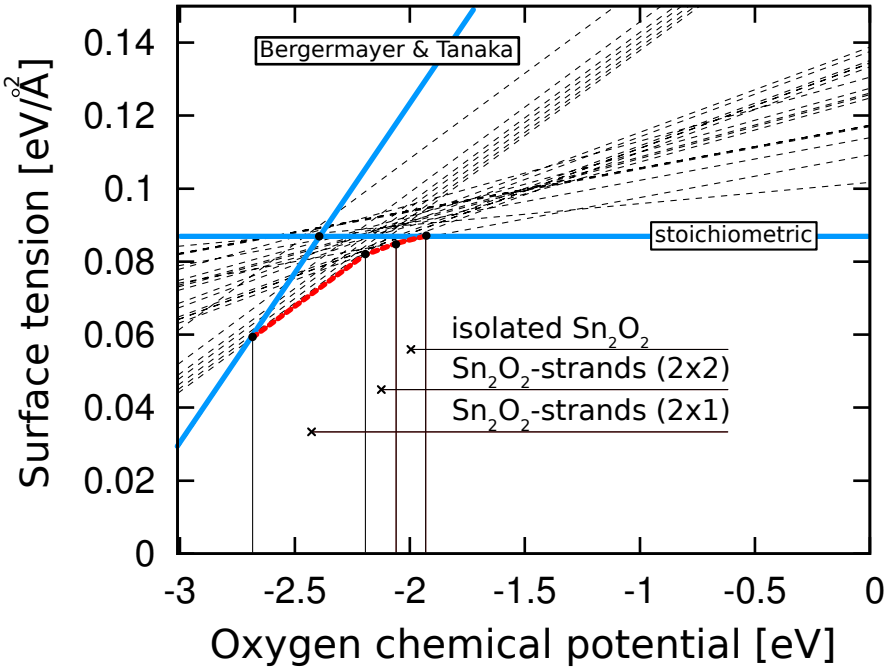


Figure 60: Surface tensions as a function of the oxygen chemical potentials for surfaces based on the Sn_2O_2 defect. The stoichiometric surface as well as the reduced surface from Ref. [310] are shown with thick blue lines for comparison. The new ground-state structures are shown with red dotted lines.

to cations necessarily being located on non-lattice positions. The existence of cation interstitial based surface terminations has been proposed already by Batzill *et al.* [305] in the context of the (4×1) reconstruction.

The proposed interstitial structures are also related to strands found on TiO_2-(110) surfaces [317, 318]. In the case of titania, however, the strand composition is driven by the formation of sub-oxide stoichiometries which are formed by Ti.

In summary, our proposed structure explains all experimental observations of the (1×1) high temperature reconstruction of SnO_2 and confirms its disordered nature. Moreover, we show that the ground-state of this disordered structure has a (2×1) periodicity which is also in accordance with STM studies.

Disordered (1×2) *reconstruction*

The high temperature (1×2) reconstruction is characterized by large height differences [301, 298, 51]. This was not only observed in STM but also with non-contact AFM measurements. Due to this distinctive feature, the geometry of the (1×2) surface cannot only be explained by the interstitial structures presented in the previous section.

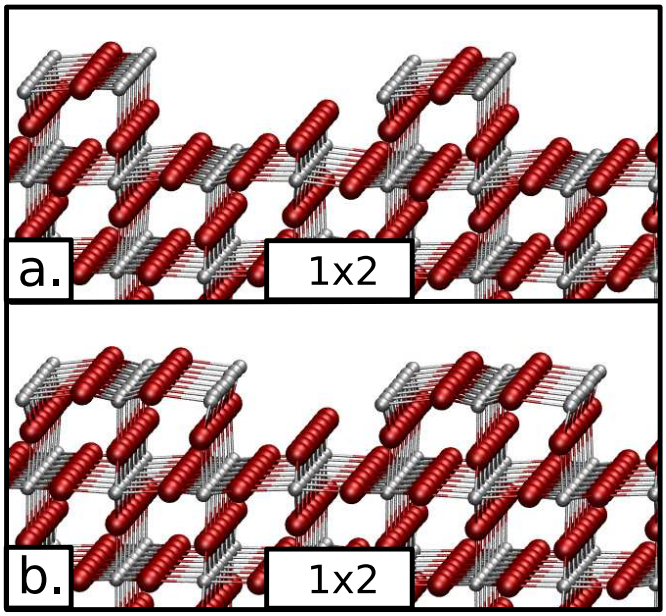


Figure 61: Basic geometries of possible (1×2) reconstructions. The bottom is showing the structure from which the more stable configurations can be derived.

In order to account for the corrugation within a (1×2) periodicity two simple and charge compensated constructions are possible and shown in Figs. 61(a. and b.).

The structure in Fig. 62(a.) is closely related to the (1×1) reconstruction proposed by Bergermayer *et al.* [310] and is obtained by mixing the stoichiometric surface with the latter alternatingly along $[1\bar{1}0]$. The structure shown in Fig. 62(b.) exhibits narrower trenches but wider plateaus. Both involve 5-fold Sn within the trenches (Fig. 61(a.)) or on the plateaus (Fig. 61(b.)) Both of these structures exhibit, however, similar and unfavorably high surface tensions which are plotted as thick dashed lines in Fig. 63 in comparison with other stable surface structures.

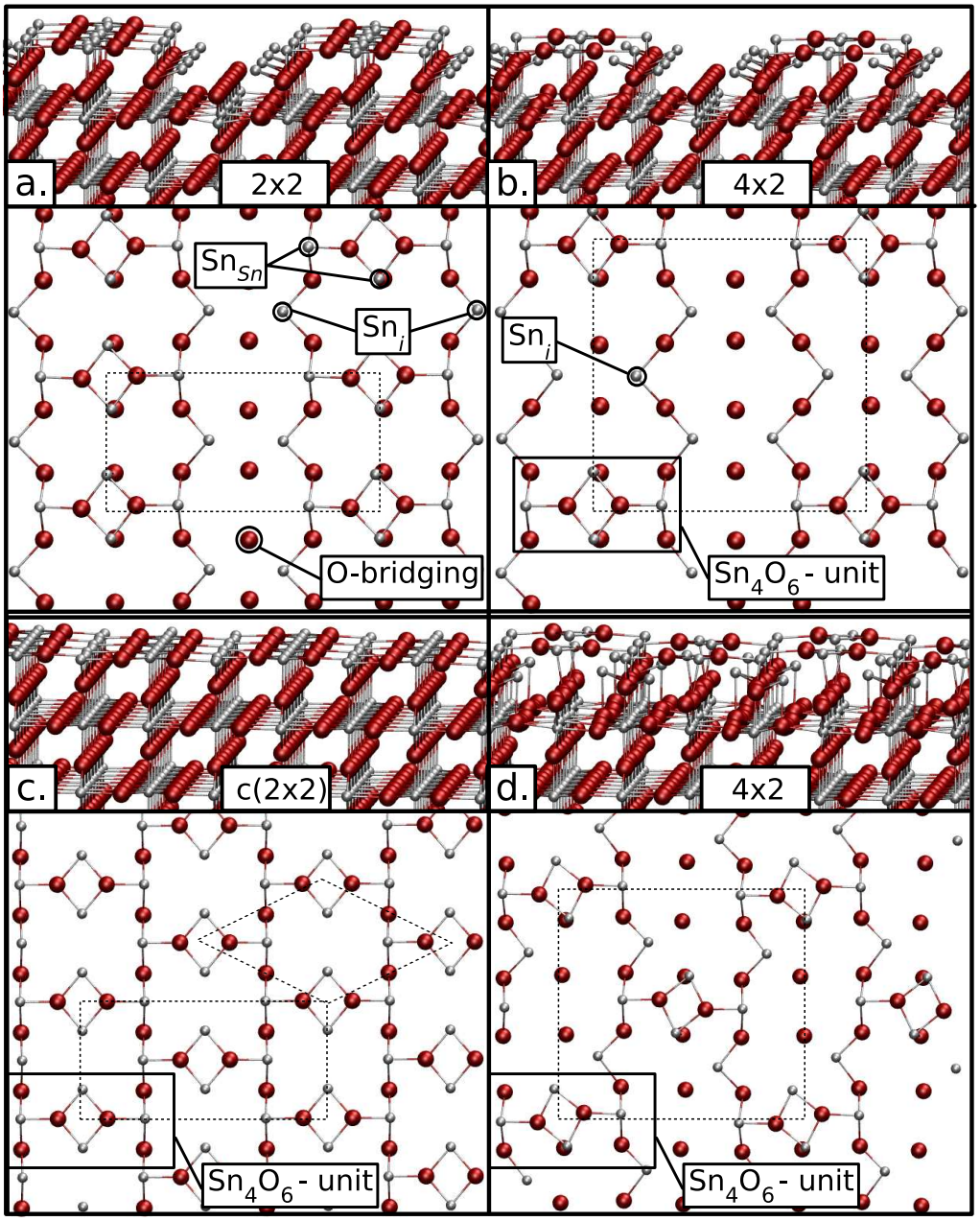


Figure 62: Geometries responsible for the (1×2) LEED pattern. (a. and b.) are thermodynamically favorable and have a different concentration of Sn_4O_6-units. The structure in (c.) is metastable and corresponds to the highest density of Sn_4O_6-units. The structure in (d.) is a junction described in the main text. Interstitial and regular lattice tin are occasionally denoted in the images.

In the following, we have combined these surfaces with the strands of Sect. 11.3.2 in the simplest possible ways (on top of Sn_{cus}) which also resulted in higher energy reconstructions with no thermodynamic stability range.

Low energy structures with a stability range can be constructed by removing every second row of oxygen along [001] from the plateaus of the structure shown in Fig. 61(b.) This procedure results in additional three-fold coordinated Sn which consequently assume the oxidation state Sn^{2+} instead of Sn^{4+}. Additionally, 2-fold coordinated tin is then located at the edges of the plateaus, which can be lowered into the interstitial positions in order to bond to a third oxygen below. This way the low energy (2×2) structure of Fig. 62 (a.) is obtained and also shown from the top. We have tested similar rearrangements also for the configuration shown in Fig. 61(a.) without obtaining surfaces with comparably low surface tension.

The thus obtained structure can also be described as a strict alternation of the geometry shown in Fig. 61(b.) plus the strands of Fig. 59 from the previous section. The new building block which is entering the disordered (1×1) reconstruction derived from the surface in Fig. 61(b.) is also framed in Fig. 62(b.) and denoted as Sn_4O_6. The difference between the structures of Fig. 62(b.) and Fig. 62(a.) is the density of this Sn_4O_6-unit. In Fig. 63 this (4×2) arrangement, where the Sn_4O_6-units are separated by three surface interstitials, shows an even larger stability range. Since both structures show a stability region in the thermodynamic equilibrium, again, we have to assume the existence of a homologous series of similar structures with comparable energy and a significant disorder along [001] which also accounts for the irregular appearance of this reconstruction in STM and nc-AFM.

It is also possible to include larger building-blocks extracted from the structure in Fig. 62(b.). Tests have, however, shown that in this case the surface tension increases rapidly with the size and no blocks larger than Sn_7O_{11} are possible *i.e.* it is generally favorable to minimize the number Sn_{cus}, which is present on the plateaus of the structure in Fig. 61(b.)

Moreover, it is possible to construct junctions which exchange the positions of the trenches and plateaus along $[1\bar{1}0]$ by one lattice vector. One calculated model for such a junction with an effective periodicity of (4×2) is shown in Fig. 62(d.). The surface tension is little higher when compared with the configurations (a.) and (b.) indicating that it is favorable to maintain continuous trenches and plateaus. The energy difference is, however, small enough to also account for the junction-like features which are observed experimentally [301].

Finally, we present the model of Fig. 62(c.) which is the configuration obtained when the density of Sn_4O_6-units is maximized. This configuration with a $c(2 \times 2)$ periodicity is slightly unstable with respect to other structures where surface interstitials are also present. This surface is highlighted in Fig. 62 and has the same stoichiometry as well as very similar geometry when compared with the model by Bergermayer and Tanaka [310].

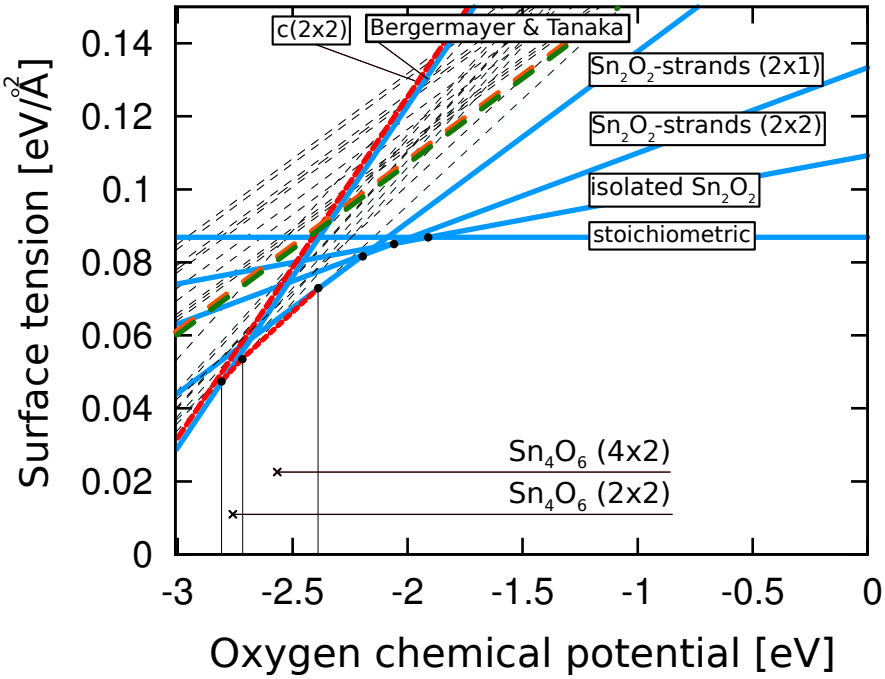


Figure 63: Surface tensions as a function of oxygen chemical potential for the disordered (1×2) reconstructions. The basic patterns from which all others are derived are dashed.

11.3.3 *Metastable structures*

According to the discussion in section 11.2, metastable surface reconstructions are relevant for the SnO_2-(110) surface. While we were able to propose new structures for the (1×1) (Sect. 11.3.2) and (1×2) (Sect. 11.3.2) reconstructions on the basis of low surface tension, there is no such simple qualifier in the case of metastable structures.

Realistic metastable reconstructions should not have exceedingly high surface tension. More importantly, small structural disturbances and little changes of the stoichiometry should result in large excess energies for true metastable structures.

Three different models were proposed in the case of (4×1), whereas we have only found one proposal in the case of the $c(2 \times 2)$. Detailed STM measurements indicate a *p2mg* symmetry for the (4×1) reconstruction [299, 300]. The atomic arrangement of the (4×1) has been conjectured to be formed by a litharge SnO-(101) coincidence lattice expanded by 10% , a model which is supported by only one other publication [301].

We have calculated models which are based on various single plane over-layers derived from the litharge-SnO lattice on top of the stoichiometric SnO_2-(110) surface. In these studies, it was always possible to find similar alternative structures which

had a periodicity different from (4×1), but with significantly lower surface tension. Therefore, we believe that such reconstructions are unlikely candidates.

As an alternative, Jones *et al.* [299] proposed an ordered surface vacancy structure. Doubts have also been raised about this model. It has been pointed out that a structure involving only surface vacancies cannot constitute a very deep local minimum in the total energy landscape [305]. Additionally, there are discrepancies concerning the corrugation as well as the STM contrast [314]. We have calculated a relatively high surface tension for this model, a surface which involves two in-plane vacancies per (4×1) unit. More importantly, we find that by associating these two in-plane vacancies, a structure is obtained with considerably lower energy, but only slight modification. Note, also that the single in-plane vacancies are no stable point defects as we have shown in Sect. 11.3.1. Thus, this surface cannot be a deep local minimum, but easily transforms into other structures. Further, due to the large number of uncompensated bridging-vacancies this surface would carry a large positive surface charge and is, therefore, not likely to be the experimentally observed (4×1) reconstruction.

Finally, we have tested the ideas by Batzill *et al.* who proposed surface interstitial configurations comparable to those presented in Sect. 11.3.2. Also in this case the resulting surface tension is large as compared to alternatives. Specifically in the case of the proposed structure, it is favorable to eject two interstitials onto the surface. Without this surplus Sn, the remaining structure converts into the disordered (1×1) reconstruction of Sect. 11.3.2.

In summary, none of the previously proposed models is in line with the experimental observations. On the other side we have not found any equilibrium candidates exhibiting the desired $c(2 \times 2)$ or (4×1) periodicity.

One model, which we came across during our studies, is however, a good candidate for a metastable structure. It can be obtained by assuming the following simplistic constraints: Stoichiometry changes occur only via oxygen removal due to its gaseous nature so that tin necessarily remains on the surface. During reduction, oxygen is removed pairwise by forming double vacancies either in bridging or in-plane geometry (lowest defect formation energies). With this assumptions it is possible to arrive at a structure with a $c(2 \times 2)$ periodicity corresponding to the one shown in Fig. 62(c). Its surface tension is also shown in Fig. 63 by a dotted red line.

The energy of this structure is very close to the one found by Bergermayer *et al.* [310] ($\Delta E \approx 0.02$ eV/$\mathrm{Sn}^{\mathrm{surf}}$). Additionally, we have calculated that any oxygen addition or redistribution of oxygen on the atomic scale leads to an energy cost of > 2 eV per process, which is high even at high temperatures. Finally, the structure is highly oxygen deficient and charge neutral (by the formation of Sn^{2+}). In reality, the formation of the $c(2 \times 2)$ is not obtained by oxygen removal but by ion bombardment which usually causes preferential sputtering of oxygen. Also, the LEED signal tends to be diffuse which indicates a significant degree of disorder. Nevertheless, this model could represent the basic structural pattern.

In the case of the (4×1) we were not able to find a similarly plausible candidate. We consider the nature of this metastable surface reconstruction to be a still unresolved issue.

11.3.4 *STM contrast*

Surfaces of SnO_2 were extensively studied by STM even with atomic resolution [299, 300, 51, 298, 301, 302, 303, 304]. We have found in the literature that generally a bright contrast is always associated with tin ions. The purpose of this section is to discuss the appropriateness of this assumption. The bright contrast for tin is generally assumed because of predominantly Sn-5*s* character of the lower CB in SnO_2. The Sn-5*s* states have an extended character which is reflected in the strong parabolic dispersion of the CB.

We note that there could be several difficulties associated with this interpretation. For example, the CB of SnO_2 is, especially at higher energies, an admixture of Sn-5*s*, O-2*p* and Sn-5*p* states. Additionally, Sn can be present on SnO_2-surfaces in several under-coordinated forms for which strong energetic variations of the Sn-5*s* position and *s*-*p* admixture may occur. These variations are most apparent by recognizing that Sn is generally able to localize a lone-pair and thereby change the oxidation state from Sn^{4+} to Sn^{2+}. For this lone-pair effect an *sp* hybrid-orbital is localized which in turn is necessarily removed from the CB. Whether still a bright feature is observed in empty-states STM for the oxidation state Sn^{2+} can, therefore, not be answered, straightforwardly. The lone-pair effect is mainly observed, whenever Sn is coordinated pyramidally either in four- or three-fold coordination, whereas the bulk-like extended character of Sn-5*s* is obtained for octahedrally coordinated Sn, which for geometric reasons cannot be located on surface sites. Realistic surface reconstructions consist of three, four, and five-fold coordinated Sn at the surface for which the appearance cannot be judged easily.

In order to illustrate these ideas, we have taken the (001) surface in the reduced and stoichiometric state and plotted the site-decomposed Sn-*s* density of states (DOS) and compared with the total DOS (Fig. 74). The stoichiometric (001) surfaces have four-fold coordinated tin at the surface (no pyramidal arrangement) beside the 6-fold coordination in the bulk. In contrast, the reduced (001) surface contains a combination of 5-fold and 3-fold tin where the 3-fold tin is in the oxidation state $2+$. The reduced surface is obtained by removing half of the surface oxygen from the topmost layer. By using the (001)-surfaces, all occurring coordinations can be covered. The DOS shows the systematic trend of the Sn-5*s* contribution in the CB as it is shifting towards lower energies by the reduction of its coordination. Note that the band gap is severely underestimated by the LDA as expected. The increasing DOS at the bottom of the CB for under-coordinated tin is, therefore, expected to lead to an increased brightness of the respective cations in STM. Neglecting topology, the brightest features should be four-fold coordinated Sn in non-pyramidal arrangement (no lone-pair localization).

The brightness of the Sn^{2+} itself is difficult to estimate but generally lower when compared with geometries without lone-pairs. In those cases tunneling can only occur into Sn-p states orthogonal to the lone-pair sp-hybrid. The energy of those states is also reduced for these lower coordinated Sn. Since the tunneling current can only enter into orbitals orthogonal to the lone-pair orbital, in the case of Sn^{2+} the brightest features may appear not exactly at the location of the ion but away from it.

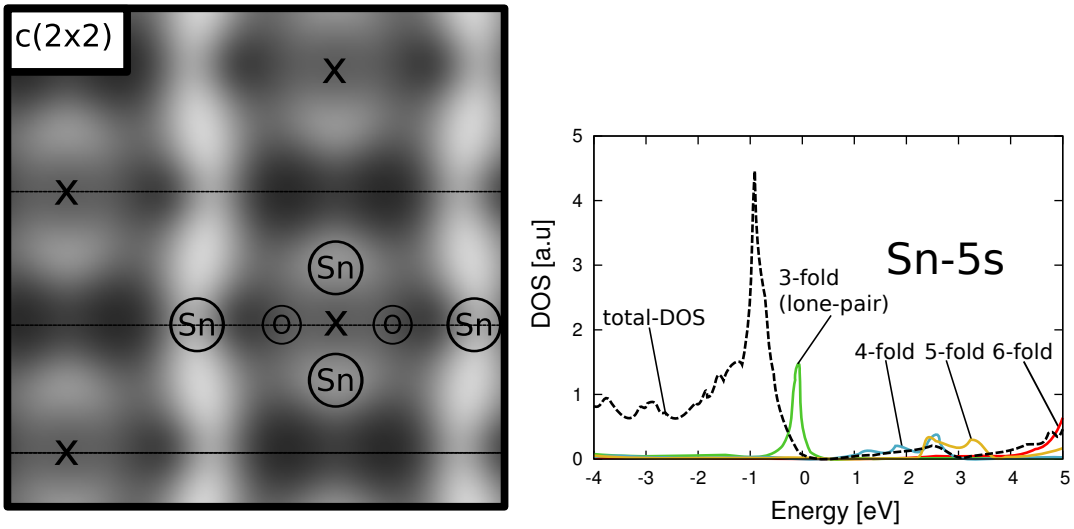


Figure 64: (top) STM simulation of a highly oxygen deficient and metastable c(2x2) surface reconstruction. The image is simulated using a bias of +1.8V. The bulk periodicity is denoted by horizontal lines whereas crosses mark the lattice sites. Also the atomic positions of the basic building block is denoted. (bottom) Electronic density of states of 3,4,5,6-fold coordinated Sn in comparison with the total DOS.

For illustration, we present a simulated STM image of the highly reduced and metastable $c(2 \times 2)$ reconstruction of Sect. 11.3.3. The image in Fig. 74 shows eight bright features per unit mesh. Also eight Sn^{2+} are present within the unit mesh and denoted in the figure (compare also with Fig. 62(c.)). It is, however, striking that only four out of eight Sn^{2+} are at the location of the brightness maxima. In fact, the brightest features in the image do not correspond to atoms at all. They are found exactly at the interstices between the other two Sn^{2+}. This observation can be crucial for the resolution of the yet unknown surface reconstructions such as (4×1).

Also, it should be noted that generally the brightness of Sn^{2+} is much weaker when compared with 6,5,4-fold coordinated tin. Therefore, if different tin species are present Sn^{2+} might not be observable at all. It is also important to note that the brightness of Sn^{2+} can vary itself depending on coordination (directionality of p-orbitals) and is sensitive to the applied bias. In our example all tin ions are at equal height showing, however, different brightness.

These ideas can also help to interpret defect structures observed in future STM studies. Finally, we note that we have also made sure that the proposed equilibrium structures of the previous sections reproduce an STM contrast consistent with experiments *i.e.* in these geometries tin has the expected bright contrast also in the oxidation sate 2+.

11.4 SUMMARY AND CONCLUSION

The SnO_2 surface has been thoroughly re-inspected with the aim of explaining the experimentally found reconstructions. We identify a class of structures whose members show similar energies and appear as equilibrium structures below an oxygen chemical potential of $\mu_O < -2.0$ eV in agreement with the observed reduction point of SnO_2-(110) surfaces [51]. Furthermore, several other experimental observations are in line with the proposed structure. Namely, this surface reconstruction is a (1×1) high temperature phase which has a tendency for disorder and shows fractional interlayer spacings. This is because its construction involves Sn on surface interstitial positions which are otherwise unfavorable species in the bulk of the Rutile phase. The overall ground-state structure corresponds to a (2×1) reconstruction with a zig-zag pattern which is also indicated in experiments.

In the case of the (1×2) reconstruction, we propose a new structure closely related to the high temperature (1×1) reconstruction. In this case, additional Sn_4O_6 building blocks are added to the strand-like features of the (1×1) structure and connected by those along $[001]$ direction in a disordered manner.

Also in this case the experimental observations closely agree with our proposed structure. The large corrugation is a basic feature of the new building block, whereas the irregularities observed in the experiments are due to the disorder and intervening sections made up of (1×1). Finally, junctions which are observed in STM and nc-AFM are energetically reasonable and compatible with this model.

Among our test-structures, we have been able to identify one plausible candidate for the metastable $c(2 \times 2)$ reconstruction. In contrast, none of the calculated (4×1) structures could convincingly be assigned to the observed (4×1) reconstruction.

In summary, we have identified thermodynamically stable models for the high-temperature (1×1) and (1×2) reconstructions of SnO_2. Based on these findings we propose a revised equilibrium phase-diagram for SnO_2-(110) surfaces (Fig.65) which we have obtained using the LDA but also the GGA-PBE functionals for comparison. While there are slight differences specific to the chosen exchange correlation functional, the new surface phases have a stability region irrespective of the functional used and in good qualitative agreement. Due to the general over and underestimation of the cohesive energies in solids for LDA and GGA respectively the energetic offset between the functionals is plausible. The true surface tensions are expected to fall between these two limiting cases.

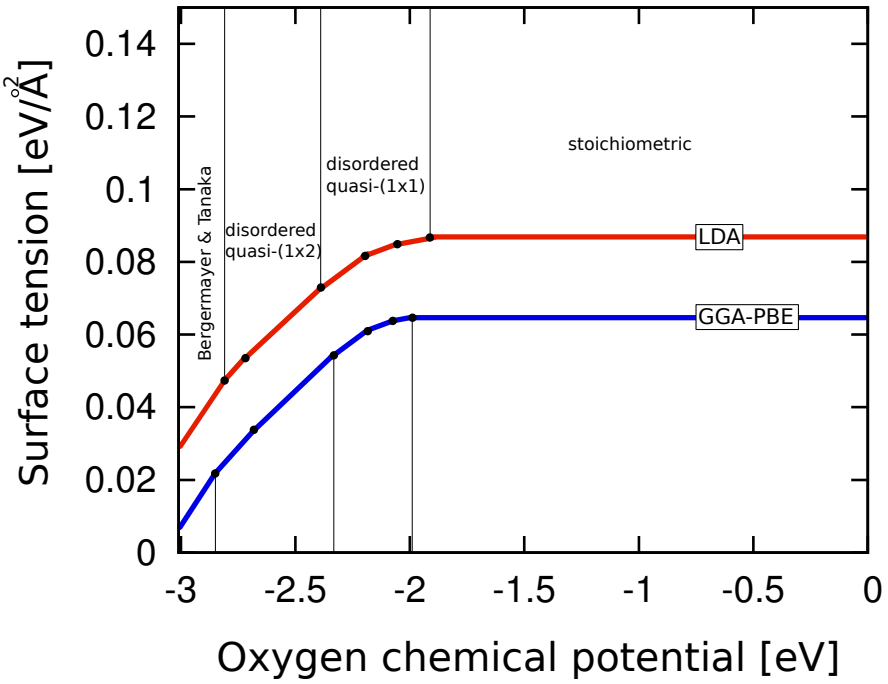


Figure 65: Revised surface phase diagram of the SnO_2-(110) surface.

We stress that according to experiments this phase diagram contains candidates for all observed equilibrium reconstructions and is, therefore, complete and consistent in this form, whereas all other observed reconstructions are metastable. The presented informations can be taken as a basis for experimental verification. It is also the starting point for an in-depth understanding of the microscopic processes in the field of gas sensing as well as interface engineering of transparent contacts.

12

SURFACES OF CUBIC INDIUM OXIDE

12.1 INTRODUCTION

The surfaces of In_2O_3 and ITO have been studied mainly by means of photoelectron spectroscopy techniques (XPS / UPS) [319, 85, 264, 262, 222, 96, 320, 95, 263] and by measuring the gas response (electrical conductivity) against various gas species [52, 53, 54, 55, 56, 57, 321]. It is found that In_2O_3 surfaces can exhibit variations in the surface composition, but also variations of the dopant concentration (Sn). These changes are often correlated with the measured surface potentials of the materials [322, 323, 85, 264, 324]. In polycrystalline samples the (111) and (001) grain orientations are often observed and the texturing appears to depend on deposition parameters, on the materials composition as well as the dopant concentration [20, 325, 326, 327, 328, 329]. It is remarkable that especially in *n*-doped samples the polar (001) surface represents the predominant surface orientation [20, 330].

In the context of the present study, experiments with heteroepitaxial films on Yttria stabilized Zirconia (YSZ) are of special interest, because they address properties of the two predominant orientations (001) and (111), directly [331, 96, 328, 332, 329, 333, 334, 335, 335, 336, 324]. Only recently, Morales and Diebold succeeded in obtaining atomically resolved images of the (111) and (001) surfaces of ITO using scanning tunneling microscopy (STM) [332, 329]. Low energy electron diffraction patterns (LEED) provide evidence for a (1×1) surface periodicity in both cases. The same authors also found that (111) surfaces remain unreconstructed [332], whereas (001) surfaces show a rich variety of surface structures, significant disorder and a dopant dependent appearance [329]. These results were interpreted on the basis of theoretical studies dealing with the thermodynamic stability [321, 337]. According to these studies the (001) surfaces may exist with a cation termination or undergo a dimerization reaction forming a surface peroxide. Several other theoretical studies were mainly concerned with the electronic structure of the ITO surfaces [338, 339, 340].

In contrast to the closely related material SnO_2 (chapter 11), and ZnO [341], the surfaces of In_2O_3 have been studied to a much lesser extent by means of electronic structure calculations. Especially, the stoichiometry variations exhibited by the surfaces as a function of the environment are not well understood. The interaction of water/hydrogen has only been studied for the (001) surfaces, but was not compared with

alternative terminations. The effect of tin has not been investigated at all, although there is an apparent influence of the dopants on the surface stability [85, 329].

In this study, we present results on the relative stability and stoichiometry variations of four experimentally observed surfaces. We study the effect of water adsorption as well as the effect of *n*-doping. We discuss the influence of lattice strain and conclude with an analysis of the electronic structure.

The methodology is mainly adopted from the previous chapter and details on the computational setup as well as the underlying slab geometries and convergence behavior can be found in Ref. [A13].

12.2 In_2O_3- SURFACES

Indium oxide crystallizes under standard conditions in the bcc-bixbyite structure, which contains 80 atoms per unit cell. A detailed description of the structure can be found in Refs. [68, 294]. The structure has an alternating anion cation stacking sequence in $\{001\}$ directions, where anion and cation layers contain 12 and 8 ions, respectively. Within the cubic In_2O_3 structure there exist two distinct cation positions, which differ by their site symmetry. We shall refer to these different indium sites as In-*b*-and In-*d* according to the Wyckoff notation.

Note that the site symmetries of In-*b* and In-*d* are S_6 and C_2, respectively. Surfaces can generally be characterized by their stacking sequence in normal direction according to the established classification scheme for ionic compounds by Tasker [213, 214]. Among the low index In_2O_3 surfaces all surface types can be found. The essential features of the stacking sequences of (001), (011), (111) and (211) surfaces are sketched in Fig. 66. More detailed illustrations can be found in later sections.

In_2O_3-(011) surfaces are classified as Tasker type-I with anions and cations in the same layer with stoichiometric composition. Two different terminations exist for the (011) orientation. Due to the existence of two different cation positions, (011) planes with only In-*d* cations and planes containing both In-*d* and In-*b* have to be distinguished. The terminations containing only In-*d* are designated

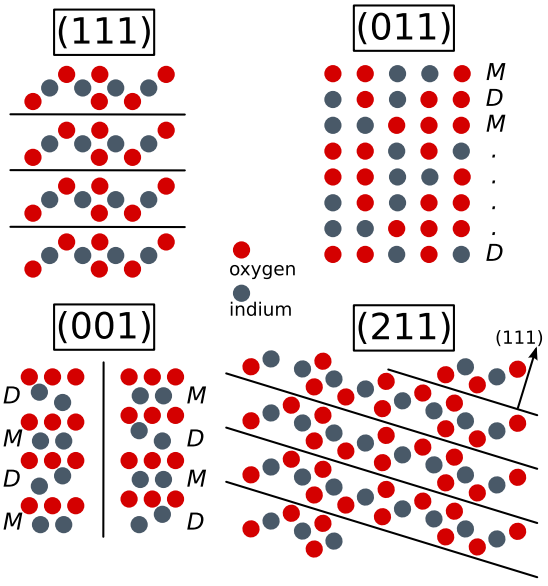


Figure 66: Sketch showing the simplified stacking sequences of low index In_2O_3-surfaces.

(011-*D*) and those containing mixed sites (011-*M*). The distribution of the anions and cations is slightly different for the two terminations as indicated in Figure 66. These two layers are stacked alternatingly and full periodicity is obtained after four repetitions of this bilayer.

In_2O_3-(111) surfaces are of Tasker type-II consisting of stoichiometric trilayers with no net dipole moment. In contrast to the related fluorite lattice, the In_2O_3-(111) surface is rather irregular and its dipole moment only vanishes when averaged over a larger area. There is one type of trilayer within the structure and full periodicity is achieved after six repetitions.

In_2O_3-(001) surfaces are polar (Tasker type-III) and consist of bilayers with a net dipole moment. This type of surfaces have generally a low stability and are unstable in the idealized model of a perfectly ionic compound. Similar to (011) surfaces, two different terminations can be distinguished [(001-*D*) and (001-*M*)] based on the content of the different cation sites. The In_2O_3-(001-*D*) cation plane exhibits some buckling, whereas the layer containing mixed sites is perfectly flat (see Fig. 66). The two bilayers are stacked alternatingly and full periodicity is reached after two repetitions of these quadlayers. It should also be noted that the bixbyite structure has inversion symmetry meaning that the polar (001) surface can be transformed into quasi type-II (unlike *e.g.* the wurtzite-(0001) surface) by means of appropriate slicing or redistribution of atoms within the surface layers.

Finally, In_2O_3-(211) surfaces have a mixed character as depicted in Fig. 66. This surface is best described by an array of short (111) facets. The step-edges of the so formed terrasses have a (100) orientation so that consequently the (211) surfaces are partially polar and can be considered as a mix of type-II and type-III. There is no evident stacking sequence for this orientation. It is, however, possible to slice the structure after each indium layer. This way, three distinct terrace structures can be constructed which make up the basic stacking units. Due to the typical bixbyite distortion these building blocks have slightly different stoichiometries.

12.3 SURFACE STABILITY

12.3.1 *Stoichiometric surfaces*

In a first, step we have constructed the stoichiometric low index surfaces (111), (011), (001), (211) and compared their surface energies (Table 14). The surface tensions increase in the following order: $\gamma(111) < \gamma(011) < \gamma(211) < \gamma(001)$. It is important to note that this ordering is changed neither by the exchange-correlation functional employed nor by the inclusion or exclusion of the In-4*d*-electrons in the calculations. The order of the surface tensions is in line with most recent calculations by Walsh and Catlow [340] as well as with the expected behavior for fluorite derived structures. The corresponding geometries are illustrated in Fig. 67. The polar surfaces have the highest surface tensions [(001),(211)]. While this is expected from a theoretical viewpoint, it is also somewhat contradictory to the very frequent experimental observation of predominantly (001) textured ITO thin-films [20].

In the case of (111) and (011) orientations, the surface structures correspond to the simplest possible bulk truncations. Both of these surfaces are built from dipole free units with bulk stoichiometry. No change of conformation is therefore needed in order to stabilize these surfaces. Consequently, only little surface relaxations are observed. The (111) surface is the overall most stable surface of bcc-In_2O_3 followed by the (011) surfaces. The $(011$-$D)$ termination is by only 1 meV/Å^2 more stable than $(011$-$M)$.

In the case of the (001) orientation, the choice of a stoichiometric surface is more complicated. One obtains surfaces with either half cation or half anion terminations. This rearrangement of charges avoids the Coulomb divergence of Tasker type-III surfaces. Since In_2O_3 contains 8 In and 12 oxygen sites within one (001) surface unit cell, the resulting stoichiometric terminations contain either 4 cations or 6 anions distributed on 8 cation or 12 anion sites, respectively. Several arrangements of half-filled cation and anion surfaces also involving both $(001-D)$ and $(001-M)$ terminations have been considered and the following is observed:

- It is more favorable to expose half-filled oxygen layers to the surface, although the difference to the lowest energy half-indium termination is not very large ($\Delta E \sim$0.02 eV/Å^2).
- For both, half-oxygen and half-indium, the energy variation for different surface atom arrangements is comparable ($\Delta E \sim 0.025$ eV/Å^2).
- For half-indium terminations it is energetically more favorable to have a $(001$-$M)$ geometry as a terminal cation layer ($\Delta E \sim 0.016$ eV/Å^2), whereas the opposite is true for half-oxygen terminations ($\Delta E \sim 0.011$ eV/Å^2)

We conclude that the stable stoichiometric In_2O_3-(001) surfaces consists mainly of half-filled oxygen planes involving a considerable disorder (because of low energetic

differences). Further, the small energetic difference between the most stable half-cation and half-anion terminations implies that areas of both, half-indium and half-oxygen, could coexist. The large density of configurational states in the case of stoichiometric In_2O_3-(001) surfaces implies a significant contribution of entropy. For a more complete modeling of the surface configuration the configurational entropy contributions to the free energy have to be included explicitly by *e.g.* the cluster-expansion method. At present we have, however, not attempted such an analysis.

The (211) surfaces of In_2O_3 are made up of terrace-like structures with mixed (111) and (001) character (Fig. 69). As it will shown below, stoichiometric surfaces of this orientation are of no particular interest as they are never stable with respect to oxygen-rich or indium-rich surfaces. For completeness, in Table 14 the value of the most stable stoichiometric variant is listed.

Table 14: Surface tensions of stoichiometric In_2O_3 surfaces calculated with different exchange correlation functionals and semi-core states included or excluded from the valence. The values are given in units of $eV/Å^2$

Surface	LDA	GGA-PBE	GGA-PBE (In-4*d*)
(111)	0.07(2)	0.05(6)	0.04(9)
(011)	0.09(4)	0.07(7)	0.06(9)
(211)	0.11(8)	0.09(5)	0.08(5)
(001)	0.14(2)	0.11(4)	0.10(2)

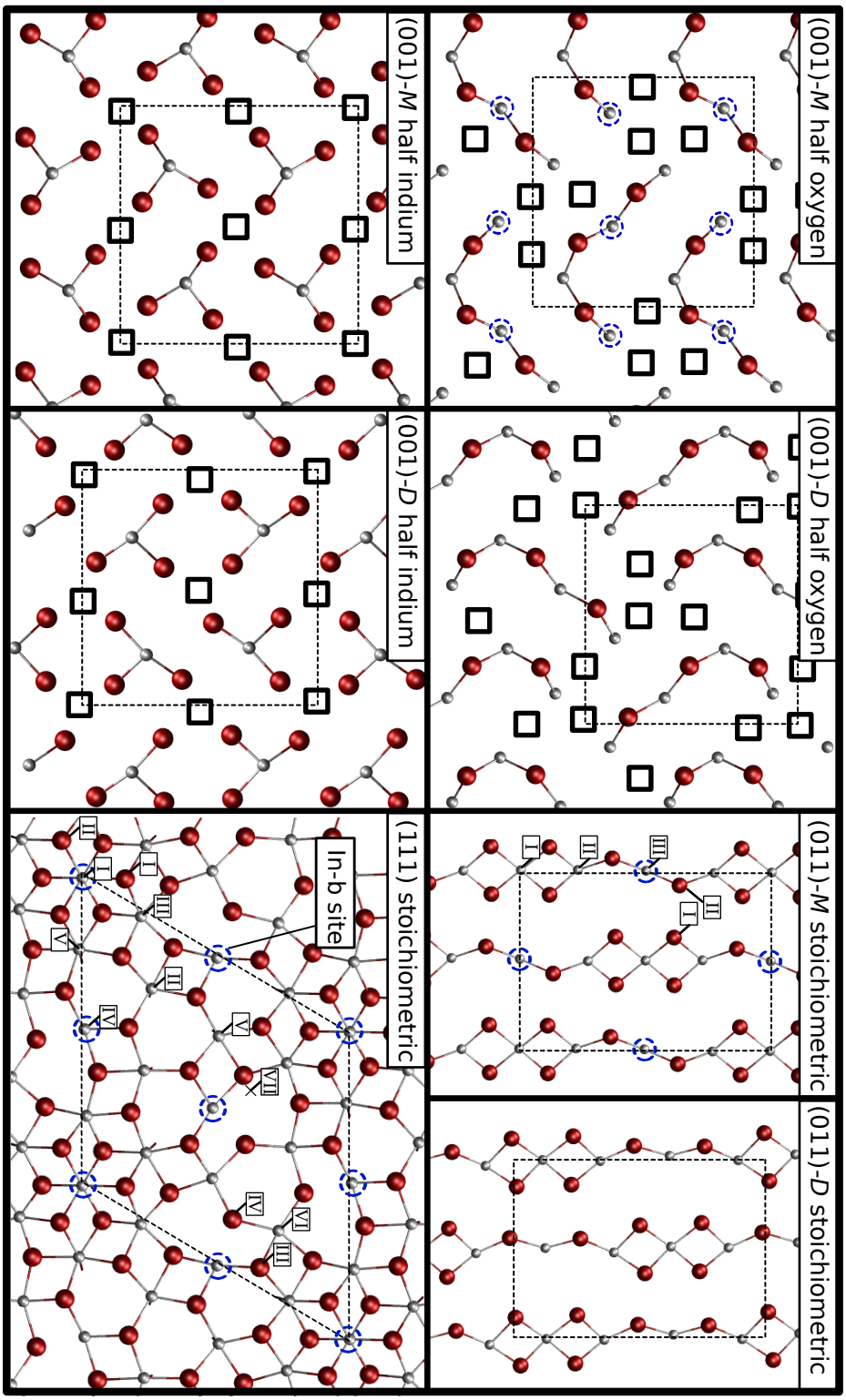


Figure 67: Representation of the stoichiometric (001-*M*/*D*), (011-*M*/*D*) and (111) surfaces. Red large and small grey balls represent Oxygen and Indium, respectively. The In-*b* atoms are circled whereas missing lattice atoms are denoted by black boxes. Numbers indicate specific cation and anion sites. Note that the sites In-I and In-V are sixfold coordinated on (111) whereas other In are fivefold. On the (011) surface In-I/II are fivefold and In-III is fourfold.

12.3.2 *Surfaces with stoichiometry variations*

Surfaces are generally able to vary their stoichiometry, if the chemical environment changes. Especially, in the case of polar surfaces compositional variations is a means for compensating polarity.

In order to find the equilibrium compositions of the different surfaces we have calculated for all surfaces a series of slabs with varying number of indium and oxygen atoms. From Eq. 4.40 (Section 4.6) we obtain surface phase-diagrams as a function of the chemical potentials of the constituents. In the case of a two component system like In_2O_3, the chemical potentials of the constituents are related by Eq. 4.7, so that only one is an independent quantity. It is convenient to use the oxygen chemical potential, which can be easily converted into the corresponding experimental environmental conditions (p_{O_2}, T) via the ideal gas law (Eq. 4.5) and thermochemical tables [180, 212].

The results of these calculations are plotted in Figures 70 a-f. The diagrams show the stability region ranging from $H_f^{In_2O_3}/3$ (oxygen-poor/indium-rich) until zero (oxygen-rich/indium-poor). These limits correspond to the decomposition reaction into the elements $In_2O_3 \rightarrow 2In + \frac{3}{2}O_2$. The oxygen-rich limit ($\mu_O = 0$ eV) is usually not achieved in contact with a (di-)oxygen atmosphere since most experiments are not conducted at oxygen partial pressures higher than one bar but at elevated temperatures. One bar of pure oxygen atmosphere corresponds to $\mu_O \sim -0.3$ eV at ambient temperature and decreases to $\mu_O \sim -0.5$ eV at modest heating of 200 °C.

Note that the oxygen-rich side of the diagrams is, however, frequently reached and probably even exceeded when exposed to less stable oxygen species like atomic oxygen or ozone (O_3). Atomic and even accelerated oxygen can be present in sputtering processes, whereas ozone or oxygen plasma is frequently used for oxidation of contaminants on electrodes [342] and improvement of the injection properties [323]. The lower end of the diagram is reached by simply heating at low oxygen pressure. The reducing limit corresponds to a temperature of 1250 K under UHV (10^{-8}Pa) conditions.

The low index surfaces of In_2O_3 show stoichiometric variations to rather different extent. Generally, the variability of the stoichiometry depends on the absolute magnitude of surface tension. The higher the surface tension the stronger the propensity towards stoichiometry variations. Larger reconstructions of the surface including stoichiometry variations can be expected when the surface energy per surface atoms is in the range of typical formation energies for point defects.

(111) *Surfaces*

The (111) surfaces exhibit a chemically inert behavior over almost the entire range of oxygen chemical potentials. We have calculated slab geometries involving different arrangements of surface vacancies by reducing this surface until all oxygen in the topmost trilayer was removed. In addition, vacancy geometries involving the topmost but also the second oxygen layer were considered. The surface composition of the (111) surface remains intact down to a chemical potential of $\mu_O \sim -3.0$ eV (Fig. 70 a).

The surface with a small stability range at low oxygen chemicatl potential corresponds to the situation with all oxygen removed from the surface located in the upper part of the terminal trilayer. This strongly reduced (111) surface exhibits mainly three-fold and one-fold coordinated indium at the surface. It is important to note that also several other reduced (111) surfaces become stable with respect to the stoichiometric surface at a very similar oxygen chemical potential (denoted by thin black lines). More generally, this point marks the oxygen chemical potential at which the surface oxygen vacancy formation energy vanishes. This behavior suggests a continuous reduction of the surface on an atomic level by the increasing concentration of surface oxygen vacancies up to a full depletion. This behavior is very similar to that found in CeO_2 [343], a structurally related material.

The oxygen surface vacancy concentration steadily increases with decreasing oxygen chemical potential up to the point where the phase transition is indicated in Fig. 70 a.). At this point the surface has reached its fully depleted state. This behavior indicates that the surface reduction is not related to a phase transition associated with a nucleation process.

Note that the vacancy formation energy considerably varies for different oxygen positions within the surface unit cell and is lowest for the oxygen denoted O-I in Fig. 67. As a consequence oxygen can be removed from these positions at considerably lower oxygen chemical potential and furthermore it implies a locally inhomogeneous vacancy distribution on the surface.

On the oxygen-rich side, it is energetically not favorable to add surplus oxygen at any site. We have considered the attachment of atomic oxygen and dimeric oxygen at various locations on the surface and always observed a positive excess energy even under oxygen-rich conditions.

(011) *Surfaces*

Chemically, (011) surfaces show stoichiometry variations very similar to (111) surfaces. Only at harsh reducing conditions, these surfaces exhibit an oxygen depletion on the surface by the occurrence of surface oxygen vacancies. The difference in comparison with (111) is that the reduction point is located at a slightly higher oxygen chemical potential ($\mu_O \sim -2.6$ eV). As for (111) also in the case of (011) surfaces the transformation between the depicted phases (Fig 70 b.) and c.)) is rather continuous. The different line segments correspond to no unique reconstruction below the reduction point and infinitely many can be constructed in between those. Nevertheless, also on the (011) surface, oxygen positions are associated with different vacancy formation energies depending on their coordinations.

The oxygen species on the stoichiometric (011) surfaces are all three-fold in geometric arrangements which vary from close to trigonal planar to trigonal pyramidal. Oxygen vacancies are preferentially produced in the latter geometry (*e.g.* O-I in Fig. 67). Because this oxygen geometry predominates in the case of (011) surfaces oxygen vacancy formation is easier in comparison with the (111) surface where a

major fraction of surface oxygen is four-fold coordinated. The reduction is also slightly different when (011-M) and (011-D) are compared. This is due to the different oxygen geometries in the two terminations. For example, on the (011-M) surface it is favorable to remove all oxygen atoms which are depicted in Fig. 67 (12 oxygen per surface unit cell) in the metal-rich limit. This surface consists only of the remaining one– and twofold coordinated In chains along $(1\bar{1}0)$ direction and is represented by the lowest line segment in Fig. 70 b ($\mu_O < -3$ eV). In comparison, the lowest line segment in the case of the (011-D) termination has still 4/12 oxygen per surface unit cell remaining and a full oxygen depletion is not favorable under any condition.

As in the case of the (111) surface, on the oxygen-rich side we could not identify any stable structures by the attachment of atomic or molecular oxygen.

(001) *Surfaces*

The In_2O_3-(001) surfaces exhibit the most significant variations of stoichiometry due to their comparatively large surface tension. We have adopted the ideas of Golovanov *et al.*[321] and Zhou *et al.* [337] and compared their results with several alternative structures in order to obtain the phase diagrams which are shown in Fig. 70 d.) and e.).

Under reducing conditions it is most favorable to simply remove all surface oxygen leaving behind a complete metal cation termination. This composition has a stability range of ~ 1 eV in the diagrams of both variants, (001-M) and (001-D).

The cation arrangement of the two terminations is illustrated in the top part of Fig. 68. While the indium atoms in the (001-M) terminal layer are all 3-fold coordinated, in the case of (001-D) it is a composition of 2,3,4-fold coordinated In. This variation of coordination also leads to a variation of the position of In atoms normal to the surface. The energetic difference between the two terminations is negligibly small in the reduced state.

At the lowest oxygen chemical potential, additional small line segments can be found which correspond to even stronger reduced cation terminated surfaces with additional surface indium on top (not illustrated). This can be thought of as a nucleation of metallic tetragonal body centered In and becomes favorable even before the stability limit is reached.

At high oxygen chemical potential oxygen-rich surfaces occur, a feature which is not found for (111) and (011) surfaces. This is particularly interesting considering that In_2O_3, but also other typical n-type TCOs like ZnO, SnO_2, or Ga_2O_3 are predominantly oxygen deficient and no super-oxides nor peroxides of indium are known.

The stabilization of the oxygen-rich surface terminations occurs by dimerization of oxygen in the topmost plane as it was already suggested by Golovanov *et al.*[321] This surface can be understood by the formation of a local peroxide. If we assume the oxidation state O^{1-} in the peroxo-anion configuration, the surface charge is formally reduced to half of a bulk oxygen layer, which is favorable in order to avoid the Coulomb divergence of polar surfaces.

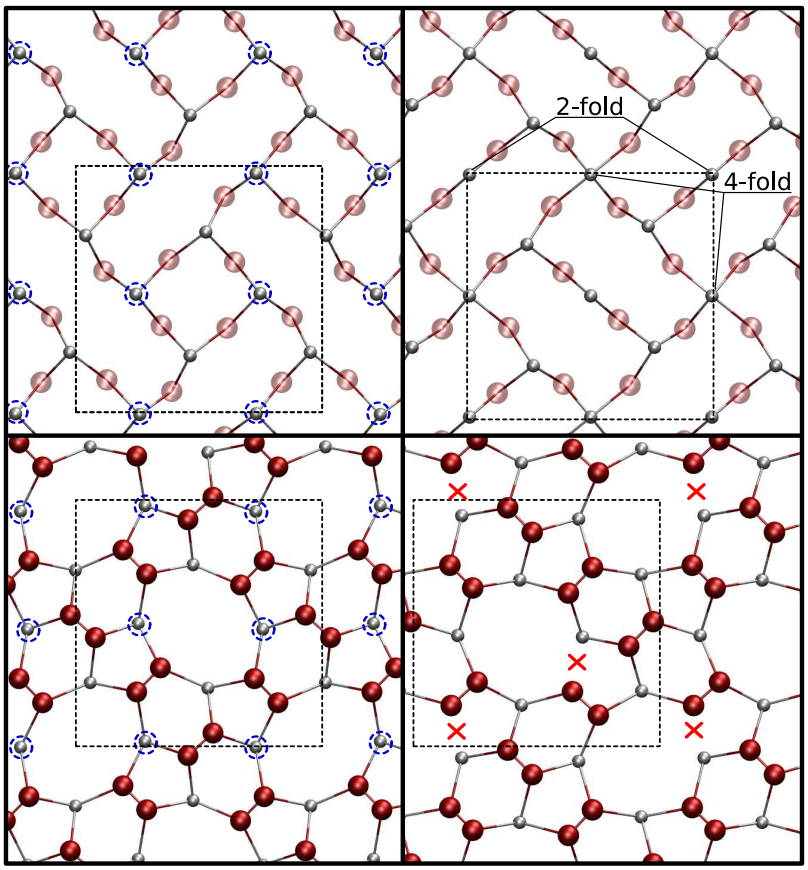


Figure 68: Representation of the metallic oxygen depleted (top) and most oxygen-rich peroxide reconstructions for (001-M) (left) and (001-D) (right), respectively. Red large and small grey balls represent oxygen and indium, respectively. A red cross indicates the location of an elongated bond showing the incomensurability of the full peroxide termination with the (001-D) cation sublayer.

The lower part of Fig. 68 shows the fully dimerized surfaces placed on the two types of indium planes. Interestingly, the fully dimerized (6 dimers per unit cell) surface is only stable on the (001-M) termination for geometric reasons.

In the case of the (001-M) surface with peroxides, all cations below the dimers attain a 6–fold coordination as it is found in the bulk. In contrast, it is not possible to arrange the dimers in a way so that the bulk coordinations are obtained in the case of (001-D). This geometrical constraint leads to a strained peroxide surface layer (see Fig. 68 (bottom right)) and a higher surface tension.

More generally, the (001-D) termination shows a tendency towards oxygen poorer surfaces. This is also manifested in the occurrence of another oxygen deficient phase between the metallic and stoichiometric which is not present in the case of (001-M).

The stabilizing feature is the highly under-coordinated (2-fold) In on (001-D), which binds oxygen anions more tightly and leads to the possibility of a partial reduction (1/3 oxygen coverage).

Between the full peroxide and stoichiometric surfaces another intermediate partially dimerized surface exists for both, (001-M) and (001-D). These surfaces consists of four oxide anions and two dimers per surface unit cell. The total negative charge of this surfaces equals that of stoichiometric and full peroxide surfaces and is therfore stable according to the Tasker classification. It is, however, noteworthy that not all possible charge compensated surfaces have a stability range.

For example, another charge compensated and partially peroxide covered surface can be constructed with four dimers and two oxide anions which, however, exhibits a considerably larger surface tension and therefore no stability range.

(211) *Surfaces*

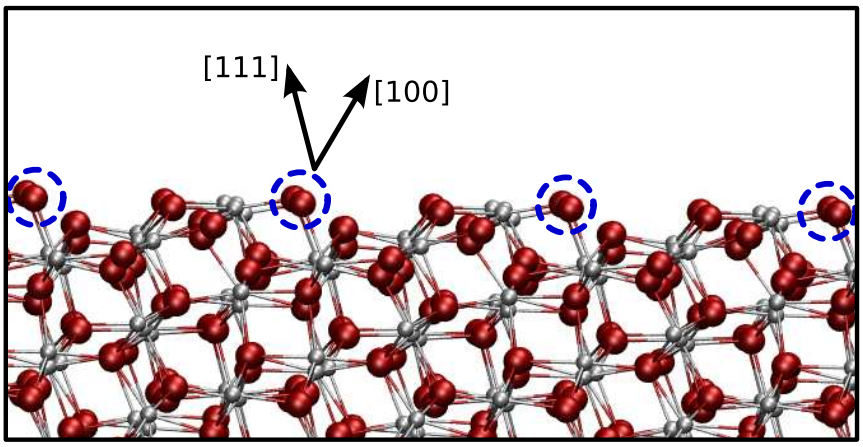


Figure 69: Representation of one stoichiometric In_2O_3-(211) truncation. Red large large and small grey balls represent oxygen and indium respectively. The circled oxygen rows are those which can be removed (reduction) or form peroxo-anions (oxidation) depending on the given oxygen chemical potential.

There are three different terminations in the case of In_2O_3-(211) surfaces. Their structure as well as their stoichiometry variations are, however, very similar. Fig. 69 shows the general structure of these surfaces in a side view. The crystallographic directions are denoted in the figure in order to indicate the relationships to the (001) and (111) faces.

The structure of this surface is a series of facets with the long edges being (111) and the step edges being the extensions of oxygen (001) planes. The step-edges themselfes have $[\bar{1}10]$ alignment in the figure. These edges are characterized by a row

of protruding oxygen which are circled in the figure. The three different terminations differ mainly by the number of oxygen atoms within these edge positions and only one cut is presented in the figure.

The stoichiometry variations of these surfaces can be understood as a combination of the observations made for (111) and (001) surfaces. While on the (111) faces no variation is energetically favorable (except for vacancy formation at low μ_{O_2}), the edges which have (001) character can be reduced and oxidized similar to the (001) surfaces. Oxygen can be removed resulting in reduced edges, whereas it is favorable to accommodate extra oxygen in the form of a peroxide at the edges. Since the edges only make up a smaller fraction of the total surface area the stoichiometry variations are weaker when compared with (001). Due to the presence of large areas with (111) character the lower surface tension in comparison with (001) surfaces can be explained. Finally, we note that the chemistry of (211) surfaces should also closely resembles that of (111) step-edges due to the structural simmilarities.

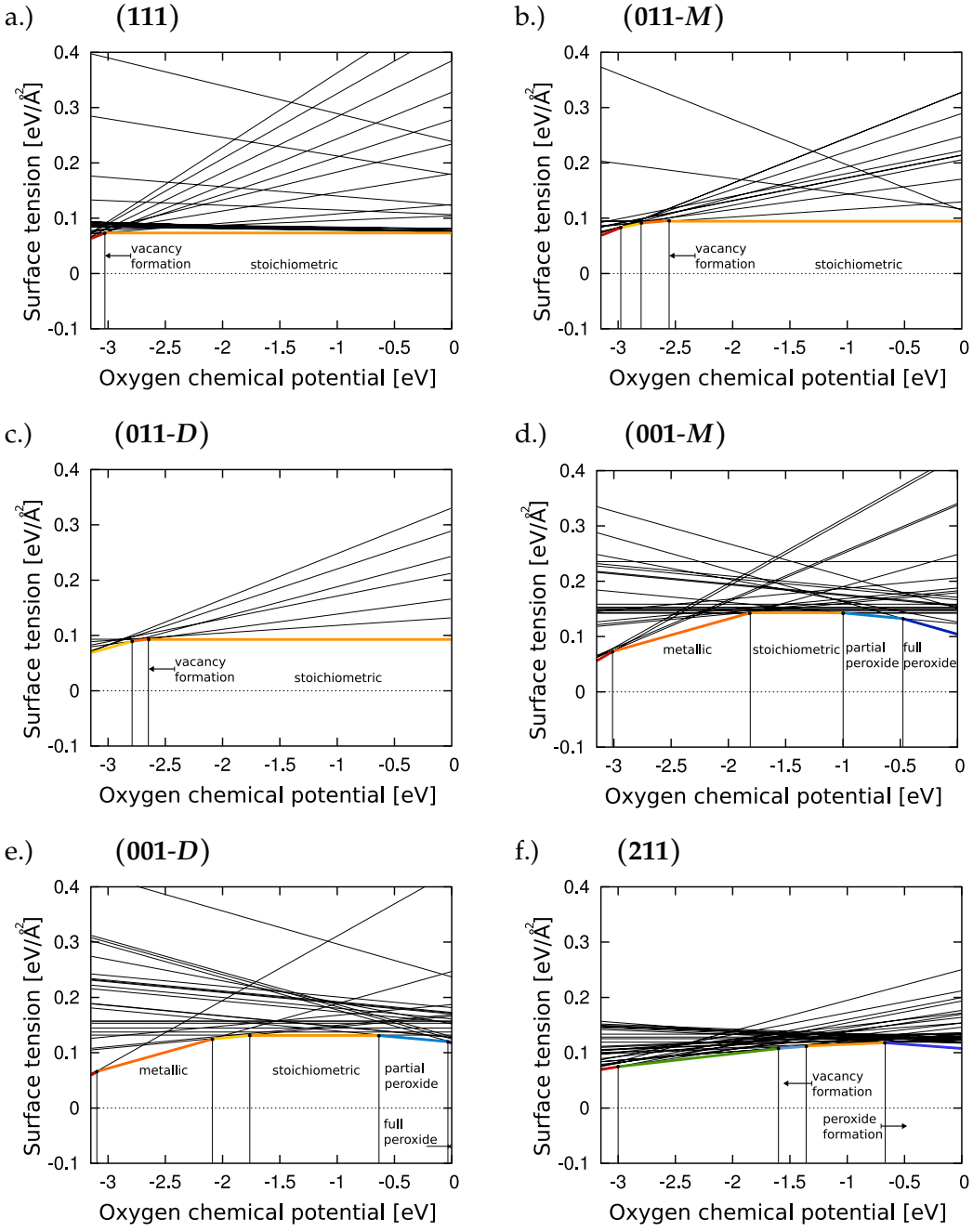


Figure 70: Calculated surface tensions as a function of the oxygen chemical potential for different surfaces of In_2O_3. Black lines denote calculations with various stoichiometries. Colored lines represent the energetically most stable surface reconstruction.

12.4 WATER

12.4.1 *Water adsorption on stoichiometric (111) and (011) surfaces*

The (111) and (011) surfaces are chemically almost inert with respect to oxygen, but it remains to be shown to what extent these surfaces can adsorb water and promote dissociation of the molecules.

On the stoichiometric (111) and (011) surfaces, species with various coordinations can be encountered. In the case of (111) surfaces, anions are either four– or threefold coordinated while cations have six– or fivefold coordination (see Fig. 67). In the case of (011-M/D), all anions are threefold coordinated in the topmost layer, whereas fivefold and fourfold Indium can be found.

In Table 15 the adsorption energies of water on various sites of these two surfaces are listed. The site numbers in this table correspond to those depicted in Fig. 67. They label the cations and anions on which the oxygen and hydrogen of the water molecule attaches to the surface.

Depending on the respective site, the adsorption energies can vary, but are in average of the same magnitude for both surfaces. The adsorption energies are in the range of $H_f^{H_2O}/2$ which makes the dissociation possible. In the case of (111) surfaces, water dissociation is mainly possible where fivefold Indium and threefold Oxygen are found in close proximity. Since such sites are only present on a small fraction of the total surface area, (111) surfaces exhibit a spatial separation of regions with different chemical activity.

In the case of (011) surfaces, a similar behavior can be observed. In this case, the dissociation can only occur on 4-fold coordinated In. The separation of sites where dissociation is observed is not so large when compared with the (111) surface. The surface unit cell of stoichiometric (011) is a centered tetragonal cell of less than half the size of the hexagonal (111) unit cell.

Further, it is found that the separation of the two hydroxyls which are created upon water dissociation tend to stay bound. For example, the energy excess due to the separation of two vincinal hydroxyls on (111) surfaces is $+0.68$ eV.

From the adsorption energies the desorption conditions for water can be inferred for the two surfaces. We estimate the desorption temperatures as a function of the water gas partial pressure by the condition [180, 212]

$$\Delta G_{\text{ads}}^{H_O}(T, p_{H_O}) = 0 \tag{12.1}$$

i.e. the conditions at which the free energy of adsorption vanishes. Considering the most strongly binding sites they are $T = 700\ K$ and $T = 350\ K$ at 1 atm and UHV respectively for (111). In the case of (011) surfaces these values are somewhat smaller with $T = 640\ K$ and $T = 290\ K$, respectively.

Table 15: List of adsorption energies (in eV) for water on stoichiometric (111) and (011) surfaces. The Indium and oxygen atoms are given to which the oxygen and hydrogen of the water molecule attach on the surface. The sites are marked in the Figure 67. The labels d. and n.d. indicate whether it is a dissociated or non-dissociated geometry of the water molecule.

(111)		(011 − M)	
$E_{ads}^{H_2O}$[eV]	position	$E_{ads}^{H_2O}$[eV]	position
-0.74	In_{III}/O_{II} (d.)	-1.02	In_{I} (n.d.)
-1.42	In_{III}/O_{I} (d.)	-0.97	In_{II} (n.d.)
-1.13	In_{VI}/O_{III} (d.)	-1.03	In_{III} (n.d.)
-1.18	In_{III} (n.d.)	-0.81	$O_{I/II}$ (n.d.)
-0.95	In_{VI} (n.d.)	-1.26	In_{III}/O_{III} (d.)
-0.70	In_{VI} (n.d.)	-	-
-0.41	O_{I} (n.d.)	-	-
-0.71	O_{IV} (n.d.)	-	-
-1.00	In_{IV} (n.d.)	-	-

12.4.2 *Hydrogenated (001) termination*

Since the non-reconstructed (001) surface has a high energy, reconstructions and variations of surface composition were necessary for its stabilization. If now water or hydrogen is brought into contact with this surface, these species can react and stabilize the surface and compete with those alternative mechanisms, like the peroxide formation. For the (001) surface and similarly for the (211) orientation it is, therefore, important to clarify the conditions under which the different compensation mechanisms are operational.

We have calculated the surface tension of the fully hydrogenated and oxygen-rich In_2O_3-(001) surfaces as a function of the chemical potentials of hydrogen and water. These surfaces are obtained by dissociatively adding one monolayer of water to the stoichiometric surface or dissociating a hydrogen molecule on each dimer of the oxygen-rich (peroxide) surfaces. The resulting surface contains twelve hydroxyls per surface unit cell. The surface tensions are plotted in Fig. 71 for different hydrogen chemical potentials (diagonal isolines). The width of the bars shown correspond to the difference in surface tension between the hydrogenated (001-M) (top edge) and (001-D) (lower edge). Overlaid in the same figure are the phase diagrams of (001-M) and (001-D) as well as the (111) surfaces for the hydrogen free case.

Evidently, the adsorption energy with respect to the isolated hydrogen molecule and the oxygen-rich surfaces is very high as can be seen from the difference of the last line segment of the In_2O_3-(001) phase diagrams (full peroxide termination)

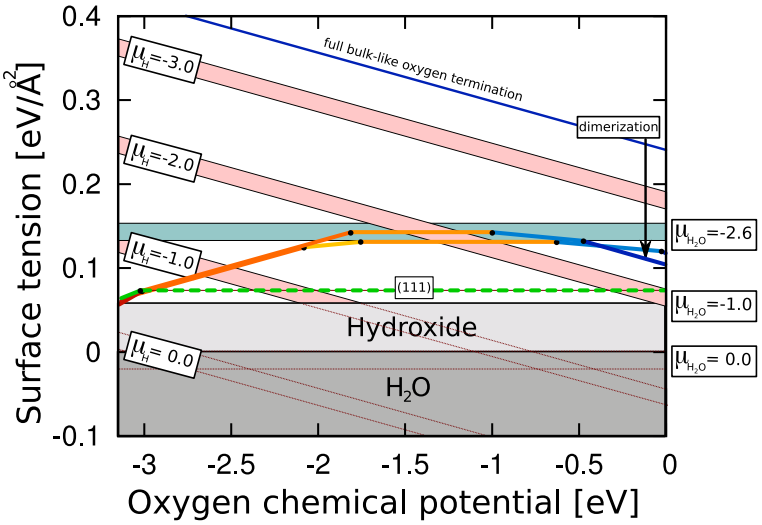


Figure 71: Phase diagram of the (001) surfaces including the fully hydrogenated surfaces represented by bars at different isovalues of hydrogen and water chemical potentials. The width of the bars corresponds to the surface energy difference between (001-M) and (001-D). The chemical potentials are given with respect to the total energy of the isolated respective molecules. The thermodynamically limiting surface tension is given at the oxide/hydroxide boundary ($\mu_{\mathrm{H_O}} = -1.01$ eV) which is indicated as the shaded area.

and the iso-line at $\mu_H = 0.0$ eV. The adsorption energy amounts to ~ -4.5 eV per hydrogen molecule. The considerably larger adsorption energy of hydrogen compared with the formation enthalpy of a water molecule warrants the dissociative adsorption process ($H^f_{\mathrm{H_O},g}[\mathrm{Exp}] = -2.50$ eV and $H^f_{\mathrm{H_O}}[\mathrm{LDA}] = -3.12$ eV). Starting with a stoichiometric surface and adsorbing water molecules instead, still a large amount of energy (~ -2.5 eV per water molecule) is gained. This value should be compared with those for non-polar surfaces (Table 15).

In fact, the occurrence of a negative surface tension in a wide range of the oxygen and hydrogen chemical potential indicates the destabilization of the material towards the formation of competing phases, namely water and the hydroxide ($In(OH)_3$). Consequently, the phase-boundaries towards these competing phases have to be considered as extrema for the chemical potentials of water or hydrogen in order to obtain the maximum stability of the (001) surfaces but at the same time maintain the thermodynamic stability of the bcc-In_2O_3.

As competing phases we have considered gaseous water and indium hydroxide in ReO_3 modification [344, 345]. The heat of formation was obtained within the

local density approximation to be consistent with the surface phase diagram. The corresponding thermodynamic constraints are

$$\Delta 2\mu_{H_2} + \Delta\mu_O \leq \Delta H_f^{H_O} \tag{12.2}$$

and

$$\Delta\mu_{In} + 3\Delta\mu_O + 3\Delta\mu_{H_2} \leq \Delta H_f^{In(OH)_3}. \tag{12.3}$$

The boundary lines of the phase transitions are given in terms of a water chemical potential, which is counted with respect to the energy of a cold and isolated water molecule. The phase transition towards the hydroxide is calculated at $\Delta\mu_{H_O} \sim -1$ eV and given by the reaction enthalpy of the water addition reaction

$$2In_2O_3 + 3H_O \rightleftharpoons 2In(OH)_3 \qquad \Delta H_r = -3.1 \text{ eV} \tag{12.4}$$

assuming that the temperature dependencies of the free energy of the solids are weak or cancel each other. Note, that the boundary is not along a constant hydrogen potential but along constant water potential and shown in Fig. 71 as the boundary to the shaded area. For chemical potentials of water/hydrogen lower than this boundary it is more favorable for the material to convert into the hydroxide.

The stability of the hydrogenated surface termination can now be compared with the (intrinsic) terminations described in section 12.3.2 along the hydroxide/oxide boundary line. Notably, the hydrogenated surface is the most stable compared with all other phases of (001) and energetically very close and even lower than the (111) surface also shown in Fig. 71. Note that the estimated correction to the In_2O_3-(111) surface tension due to hydroxylation is small (~ 0.01 eV/) and therefore considered in the figure.

In order to fully remove water from the surface in equilibrium, its chemical potential has to be set to $\mu_{H_O} \leq -2.6$ eV. This water potential corresponds to a temperature of $T \sim 1200$ K at 1 atm and ~ 580 K under UHV. The high equilibrium desorption temperatures of water indicate the facile and persistent hydroxylation of (001) surfaces which is very distinct from the behavior of the non-polar surfaces.

We stress that the desorption of water but also the hydroxide formation could be prevented by the presence of kinetic barriers so that all values are to be understood as equilibrium estimates. Ideal growth condition of films with a (001) orientation are given at a water chemical potential at or little beyond the phase boundary towards the hydroxide ($\mu_{H_O} \sim -1$ eV).

Finally, we remark that besides the hydroxylated surface we have also tested slabs with hydride terminations (negatively charged hydrogen on top of a cation layer). Surfaces of this kind are plausible from an electrostatic point of view and hydride ion incorporation in the form of point defects has already been proposed for In_2O_3. We find, however, that this type of surface termination is unstable even at hydrogen-rich and oxygen-poor conditions.

12.5 SURFACE STRESS

In the past, a significant number of studies have dealt with epitaxial In_2O_3 films [324, 331, 96, 328, 332, 329, 333, 334, 335, 335, 336]. Especially, the morphology of (001) surfaces is rather diverse and depends on the respective growth conditions, doping level and film thickness. For example Bierwagen *et al.* [336] have reported that the most homogeneous films can be obtained at rather reducing conditions, whereas *n*-type doping generally improves the homogeneity. As a substrate material yttrium-stabilized zirconia (YSZ) is most frequently used which introduces a tensile stress in the In_2O_3 films [328]. When cracking occurs due to the in-plane stress, cleavage propagates along $\langle 011 \rangle$-directions which is plausible because (111) surfaces have the lowest surface energy.

In this section, we further explore the influence of in-plain strain $\epsilon_{\text{in-plane}}$ on the surface stability and composition for the case of (001) surfaces. The reason for these studies is the observation of a relatively large surface stress within calculations of oxygen-rich (001) (peroxide-)surfaces. This finding is in line with the non-commensurability of the peroxide termination noted already in section 12.3.2. The dimers exert a compressive stress at the surface in order to maximize their attractive electrostatic interaction to the layer below. For the same reason the dimers of the peroxide termination show a pronounced inward relaxation lowering their position normal to the surface in comparison with the oxide anions. By virtue of this observation the surface in-plane strain is expected to have a more pronounced effect on the peroxide surfaces when compared with other surface phases of this orientation. The calculated surface stresses for different phases of the $(001\text{-}M)$ termination are listed in Table 16 at the calculated equilibrium lattice constant. The surface stresses follow a similar trend for LDA and GGA-PBE, despite the differing lattice constants obtained by the two approximations.

Table 16: Surface stresses for four different (001) surfaces calculated with two different XC-functionals at the calculated equilibrium lattice constant. The values are given in units of meV/Å^2.

Phase	LDA	GGA-PBE
Metallic	-93.80	-64.04
Stoichiometric	-14.56	-2.71
Peroxide	-501.45	-416.18
Hydroxylated	+22.41	+10.20

The surface stress σ_i^s is determined by two factors according to Shuttleworth [346] and can be written as

$$\sigma_s = \gamma_i^s + A\frac{\partial \gamma_i^s}{\partial A} = \gamma_i^s + \frac{\partial \gamma_i^s}{\partial \epsilon_{\text{in-plane}}}. \tag{12.5}$$

The first part is the contribution due to the surface tension γ_s while the second part describes the change of surface energy with strain. As we are only interested in the changes of the phase transitions between different phases the surface tension drops out of Eq. 12.5 because of the condition $\gamma_i = \gamma_j$ at the point of the phase transition between phases i and j. Variations of the phase transition as a function of the chemical potentials can therefore only occur due to a different strain derivative of the surface tension for two phases. This quantity corresponds to the difference in the surface stresses for two phases.

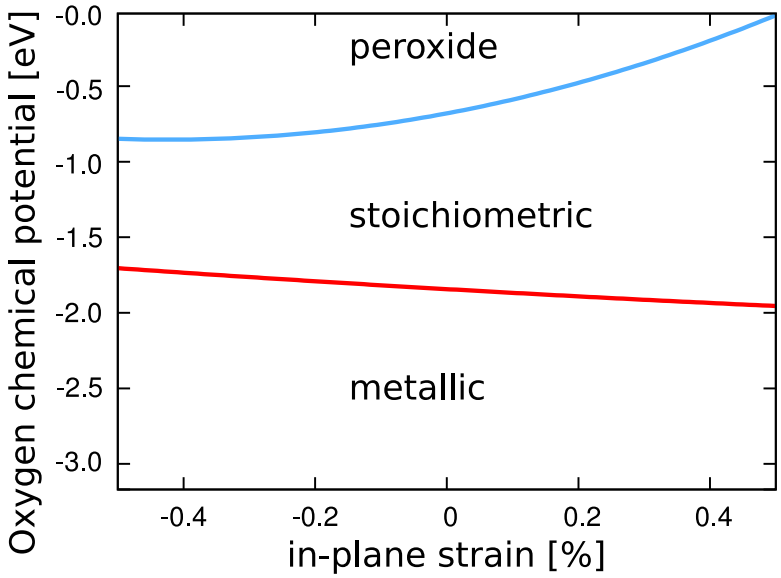


Figure 72: The strain dependent surface phase diagram for the In_2O_3-$(001\text{-}M)$ surface including only the predominant phases.

Thus, the surface phase-diagrams presented in Section 12.3.2 are generally a function of the in-plane strain (the normal direction is always assumed to be stress free) with the surface stress being the corresponding conjugated thermodynamic variable. In the case of a non-varying surface stress for different phases only a constant offset must be added to the surface tensions.

In order to illustrate the strain dependent behavior more clearly, we have recalculated the phase diagram for the $(001\text{-}M)$ surface considering only the most stable metallic, stoichiometric and full peroxide surfaces. A strain was applied in the surface plane with full ionic relaxation. The phase diagrams were then recalculated using Eq. 4.40 (section 4.6) and the thus obtained phase boundaries as a function of strain are plotted in Fig. 72.

According to the differences in surface stresses the stability of the dimerized surface is diminished towards tension while it increases under compression. More generally it can be stated that under tension, which is the case for growth on YSZ, the

stability region of the stoichiometric surface is widened. Finally, we note that for the hydroxylated surface this strong strain dependence is not observed. As a consequence, hydroxylation of films is even more likely in competition with peroxide formation under tension.

These findings have also implications in the context of gas sensor applications. It can be assumed that the strain dependent phase transitions can also lead to an either nano-particle size or film thickness dependent gas response. It is also assumed that the findings can partly be transfered to (211) surfaces as well as (111) (the edges of (111) have (001) character).

12.6 DOPING EFFECT

12.6.1 *Phase stability*

Since in most technological applications Sn dopants are added to In_2O_3, understanding the role of *n*-doping in the context of surface structure and stability is of major interest. When dopants are introduced into the material, two distinct effects have to be considered separately.

First, due to the presence of dopants the Fermi energy of the material is changed. This effect is especially pronounced in ITO where variations of the Fermi energy of $\Delta E_F > 1$ eV are possible. For this effect it is, however, important to recognize that the Fermi energy has no or only a minor influence on the energetics of charge balanced surface reconstructions *i.e.* the surfaces which are usually stable. This is because the electronic states created at surfaces of more ionic compounds are to a large degree resonant with the host states of the material and cannot be filled or depleted for Fermi energies away from the band edges. As a result, the phase transitions between the stable surfaces presented in section 12.3.2 are not expected to change because they obey electron counting rules and cannot accomodate extra charge at the surface (the metallic surface is an exception). Only the stability of charge unbalanced surfaces, which introduce additional states in the band-gap region can be affected, significantly.

A second effect is encountered when dopants are located at surface sites and directly interact with the surface species or change their electronic configuration and energy. It is possible, for example, that the oxidation state of a donor changes from the bulk to the surface. Such effects have been observed in the case of SnO_2, where the Sb dopants exhibit a lone-pair effect when located at the surface [347].

In order to disentangle these two effects, we have conducted a series of slab calculations with Sn donors at two different concentrations and with two different locations in the vicinity of $(001 - M)$ surfaces. For these studies we have included the ground state terminations, which we have found for the pure material (metallic, stoichiometric, partial peroxide and full peroxide terminatios in Fig. 70). As stated, only a minor influence of the Fermi energy is expected for these surfaces and therefore we have

additionally included another three charge unbalanced surface terminations in order to illustrate the effect of *n*-type doping.

One of these surfaces exhibits oxygen deficiency (4/12 oxygen per surface unit cell), whereas another has a surplus of oxygen (8/12 oxygen per surface unit cell). The third candidate is the peroxide surface of Sect. 12.3.2 with two dimers split into oxide ions (4 dimers and 4 oxygen per surface unit cell).

At this point, we focus on the (001) surface since only for this orientation, stoichiometry variations appear to be relevant. We believe that the principle findings, can be transfered to the (211) surface as it has some similarities with (001) surfaces (see Sect. 12.2)

The calculated relative surface tensions with Sn are shown in the four panels of Fig. 73 (thick lines). The respective slab geometries are schematically indicated in the insets. In addition, all panels contain the reference phase diagram of *pure* In_2O_3 (thin black lines). All diagrams are energetically aligned to the corresponding stoichiometric (horizontal line segment) surfaces. This allows to study variations of the phase boundaries due to the presence of dopants. The surface tension is now given as a relative value per surface unit cell. Note, that the lines corresponding to the charge unbalanced surfaces (not stable without dopants) are highlighted in red and significantly influenced by the dopants. The particular changes of individual surfaces are also indicated in the figures by arrows.

The two panels at the top of Fig. 73 show the result for tin dopants placed into the first (left) and the second (right) subsurface layer at a surface concentration of $c_A \sim 2\ \mathrm{nm}^{-2}$. With the Sn atoms in the second layer the compositional domains for the phases (blue thick lines) hardly change, whereas the phase boundaries do shift when the Sn is located within the first surface layer. Evidently, tin atoms destabilize the dimerized surfaces and reduce their stability width compared to pure In_2O_3. This is denoted as *direct effect* in the figure. The behavior can be rationalized in terms of different electronegativity and size between Sn and In. Peroxides are generally stabilized by elements with low electronegativity like in *e.g.* Na_2O_2 or BaO_2. Neither indium peroxide nor tin peroxide exist as solid phase. The stability of tin peroxide, however, if it existed, would be lower on the basis of its higher electronegativity in comparison with In. Further, due to the size difference of the cations the dimers slightly distort making the peroxide coverage even less commensurate than in the pure In case (see discussion in Sect. 72). In this context, it is also interesting to note that SnO_2 surfaces have already been studied in considerable detail and no peroxide based surface termination could be identified by experiment and theory, which provides support for the above mentioned arguments [51]. Furthermore, peroxide type point defects are less stable in SnO_2 when compared with In_2O_3 [348].

Besides this *direct* effect, large changes in both phase diagrams can be identified for the charge unbalanced terminations (red thick lines). The changes also coincide quantitatively for Sn in the first and second layer indicating the effect of only the Fermi energy. The changes are more pronounced for the oxygen-rich phases.

Going to higher doping concentrations (bottom part of Fig. 73) all effects mentioned before are enhanced and can lead to new thermodynamically stable surface reconstructions (red lines appear below blue lines). The fact that these new surface phases appear more dominantly with tin in the first layer indicates also a segregation tendency of tin to the surface as described in more detail in Sect. 12.6.2.

The donor dopant stabilized surfaces are generally characterized by a larger content of negative charge at the surface. This is plausible since the donor provides the electrons for the formation of the anions from neutral species and additionally exerts an electrostatic attraction. This can be achieved in different ways:

Starting from the full peroxide covered surface this is possible by the splitting of the dimers into oxide anions

$$O^{2-}_{2,\mathrm{surf}} + 2e^- \rightleftharpoons 2O^{2-}_{\mathrm{surf}}, \tag{12.6}$$

Starting with a stoichiometric surface this is achieved by additional attachment of oxide anions

$$O^{\mathrm{gas}}_2 + 4e^- \rightleftharpoons 2O^{2-}_{\mathrm{surf}} \tag{12.7}$$

and only to a lesser extent by the attachment of additional peroxide anions

$$O^{\mathrm{gas}}_2 + 2e^- \rightleftharpoons 2O^{2-}_{2,\mathrm{surf}}, \tag{12.8}$$

because they are destabilized by the tin dopants in the first surface layer. Additional peroxide attachment to fully peroxide covered surfaces is not favorable because of the low electrostatic interaction between the peroxide molecules and the surface cations in this case. The attachment of additional peroxide anions can only occur if the molecules are not flat but oriented perpendicular to the surface for geometrical reasons.

An interesting analogy is that the equilibria described by these reactions equivalently occurs in the bulk of the material as well. Equation (12.7) corresponds to the defect model proposed by Frank and Koestlin [109] and the peroxide defects have been described in recent DFT defect calculations [258].

At the surface this mechanism is also equivalent to the so called ionosorption model on semiconductors [349, 350]. Within this model the surface terminations of ITO are obtained by ionosorbing additional oxide-anions to the stoichiometric surface of pure In_2O_3.

We remark that the actual ionosorbed species are O^{2-} and not O^-. This result was verified by explicitly checking the oxidation state and magnetic moment of oxygen on such surfaces.

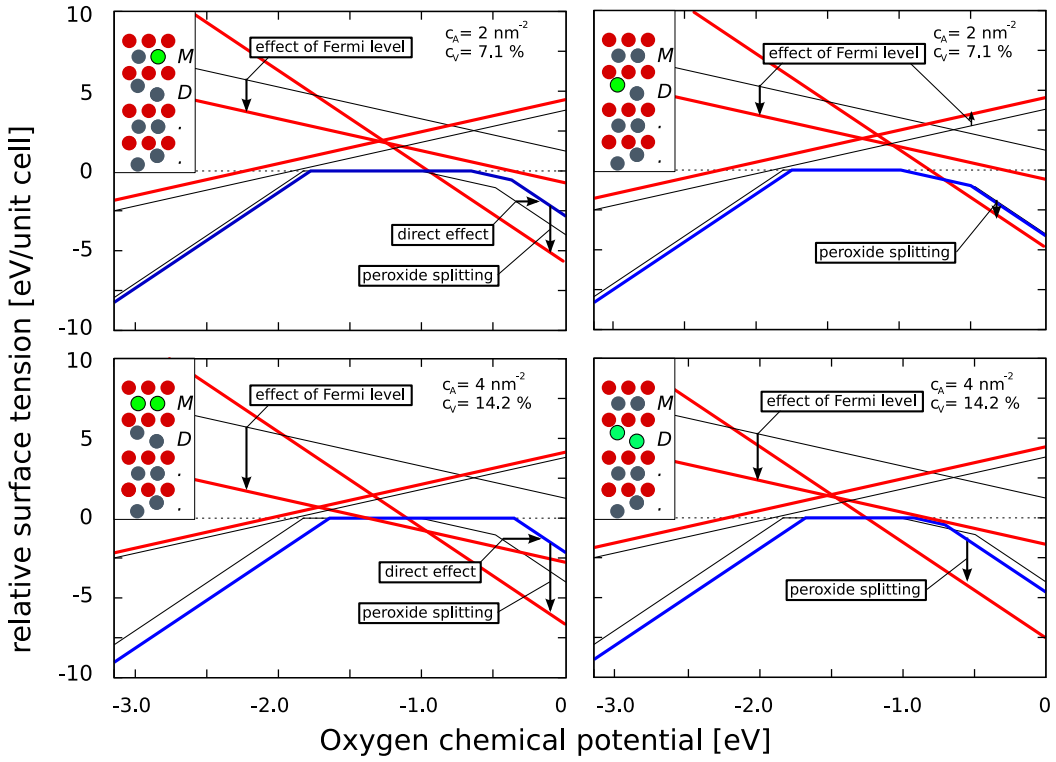


Figure 73: Phase diagrams of the $(001\text{-}M)$ surface with a tin surface concentration of $c_A \sim 2$ nm^{-2} (top) and $c_A \sim 4$ nm^{-2} (bottom) with the dopants located in the first (left) and second (right) cation plane. The setup is schematically illustrated for each case. The green atoms represent the dopants in the insets. Thin black lines correspond to surfaces of pure In_2O_3, whereas thick lines show the surfaces with Sn dopants. The charge unbalanced surfaces (unstable for pure In_2O_3) are highlighted in red. All surface tensions are normalized to their respective stoichiometric surfaces (horizontal lines).

12.6.2 *Segregation*

In the previous section the absolute values of surface tensions were abandoned, in order to study the effect of dopants on the phase transitions. However, also the absolute surface tensions vary with the concentration of dopants located within the surface. In the case of ITO, where large contents of dopants are used, their segregation can lead to a significantly reduced surface tension, which is also observed in experiments [262].

The segregation energy is defined as the energy difference between a dopant in the surface and in the bulk

$$E_{\text{seg}} = E_{\text{Sn-surf}} - E_{\text{Sn-bulk}}. \tag{12.9}$$

and can be understood as the gain or loss of surface tension per dopant atom. Table 17 summarizes the calculated values for various surface orientations and also different surface compositions. For (111) and (011) surfaces the segregation energy is predominantly positive. In the case of (111), the sixfold In sites on the surface especially those at the origin in Fig. 67 have negative segregation energies, whereas for all fivefold In the values are significantly positive.

Like for water dissociation, the segregation of Sn at (111) surfaces is spatially selective and offers an interesting way of functionalizing this surface. A similar behavior is found for (011) surfaces with the segregation energies being even more positive. A slightly negative energy is found for Sn at the fivefold coordinated In-site. Generally, the segregation energy on (111) and (011) surfaces is more negative when bound to a larger number of under-coordinated oxygen. Note, that a lone-pair effect (Sn^{2+}) is not observed for Sn in these surfaces.

In contrast to these latter surfaces, large energy gain is associated with tin beneath the (001) surfaces. An energy gain of up to 0.8 eV/Sn is observed especially for oxygen-rich surfaces. Note, that the segregation energies for (001) surfaces are taken as differences between Sn in the first and second layer. The resulting segregation energy is, therefore, likely to be underestimated. The two positive values in Table 17 for (001) surfaces are always in conjunction with full or partial peroxide surfaces and consistent with the peroxide destabilizing property of tin discussed in Sect. 12.6.1.

Note, that the segregation energies for (001) surfaces are given for two different Sn surface concentrations. This is necessary because the oxidation state of the topmost oxygen anions critically depends on the availability of electrons donated by Sn. For this reason the segregation energies increase with Sn content at some surfaces. In reality the necessary electrons could be provided by the bulk of the material which in a slab calculation is severely limited in size and, therefore, not possible unless the concentration of Sn is high enough. Accordingly, oxygen-rich surfaces can capture electrons from the (*n*-type) bulk of the material and exhibit an electron depletion at the surface.

The attraction of Sn dopants to oxygen-rich (non-peroxide) surfaces can simply be understood on the basis of electrostatic arguments. Remarkably, a segregation tendency is also observed for oxygen-poor surfaces. In this case, the driving force cannot be

an ionic interaction but is the lone-pair localization and formation of Sn^{2+} instead of Sn^{4+}. This observation is of experimental interest, because the lone-pair effect is exclusively observed on oxygen deficient (001) surfaces, while it is not observed for any other orientation (except oxygen deficient (211)). This finding offers a device for the identification of reduced ITO-(001) surfaces, if it is possible to assign deep gap-states on ITO surfaces to originate from Sn-centers.

Finally, the results of the current and previous sections also offer a possible explanation, why the (001) orientation of films is more predominant in the case of ITO, but very difficult to obtain in the case of pure In_2O_3. Taking a tin surface composition of 50%, the surface tension of the (001) surfaces can go below that of the (111) surface at oxygen-rich but also for oxygen-poor conditions. In order to see this, one can add the segregation energy per Sn (of the stoichiometric surface) to the phase diagrams of Fig. 73 (the phase diagrams were previously aligned to the corresponding stoichiometric surfaces).

A surface concentration of 50 % seems relatively large, however, the electrostatic interaction is a long–ranged effect so that energy gain is also associated with Sn in deeper subsurface regions. The exact evaluation of the surface tension is difficult unless the distribution of tin and the corresponding band bending are known in the near surface region. Evidently this distribution depends on the deposition and can be altered with the oxygen partial pressure [262], which has been demonstrated also in experiments.

Table 17: Segregation energies of Sn on (111), (011) and (001) surfaces in eV/Sn. For (111) and (011) the index corresponds to Sn located on the respective site in Fig. 67. In the case of (001) the index corresponds to the phases in the following way: 1.) metallic, 2.) 1/3 oxygen 3.) partial peroxide (1/3 O, 1/6 dimers), 4.) partial peroxide (1/6 O, 1/3 dimers), 5.) stoichiometric, 6.) full peroxide, 7.) 2/3 oxygen termination. The segregation energies are given for two different tin concentrations in the case of the (001) surface.

Site	(111)	(011)	Surf.	$(001-M)$	
				$2nm^{-2}$	$4nm^{-2}$
In_I	-0.54	0.65(7)	1	-0.36	-0.55
In_{II}	0.15	0.75(4)	2	-0.42	-0.59
In_{III}	0.00	-0.15(9)	3	-0.04	-0.20
In_{IV}	0.44	-	4	0.18	-0.09
In_V	-0.19	-	5	-0.39	-0.51
In_{VI}	0.46	-	6	0.24	0.10
In_{VII}	0.19	-	7	-0.47	-0.79

12.7 ELECTRONIC STRUCTURE

12.7.1 *Band bending*

(111) and (011) surfaces

In section 12.3.2 it was shown that the surfaces of In_2O_3 can by reduced and oxidized. This is a rather different behavior in comparison to *e.g.* SnO_2, which can only be reduced. In the case of SnO_2, the reduction is possible by the formation of Sn^{2+} instead of Sn^{4+}. Also indium has a lone-pair effect which is, however, much weaker. Within the group *III* elements it is only Tl which has a dominant $+1$ oxidation state instead of $+3$ even though In(I) compounds do exist [71]. The predominant oxidation state is important in the context of surface gap-states, but also for the expected STM contrast [351].

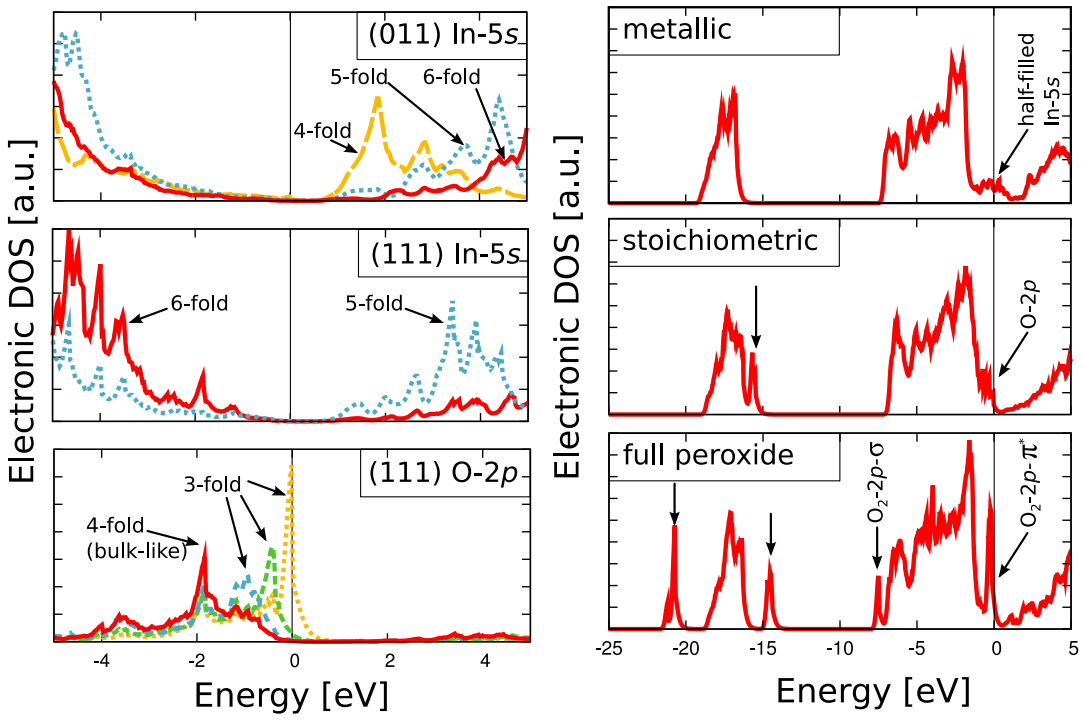


Figure 74: (left) The partial electronic DOS of In-5s states for (011) and (111) surfaces as well as the O-2*p* states of oxygen atoms on the (111) surface.
(right) Total electronic DOS for the metallic, stoichiometric and full peroxide terminations on In_2O_3-(001-*M*). The surface states are indicated with arrows. Note, that the In-4*d* semi-core states are not shown.

In the case of the predominantly stoichiometric surfaces [(111) and (011)], we have confirmed by means of Bader charge analysis the expected formal oxidation states of $+3$ and -2 for In and O, respectively also for the surface species. These surfaces can be reduced at relatively low oxygen chemical potential only and remain stoichiometric to a large extent.

Generally, due to the under-coordination of the cations on these stoichiometric surfaces the In-5*s*, states are lowered and attain a higher density of states at the bottom of the CB. In Fig. 74 on the left the site and *l*-decomposed electronic density of states (DOS) of indium cations is presented for six- and fivefold coordinated cations on the (111) surface as well as six-, five- and fourfold coordinated cations on the (011) surfaces (in this case sixfold coordination corresponds to a bulk atom). Consistently, the under-coordinated In atoms show an increasing *s*-contribution to the bottom of the CB. Since in the case of In_2O_3 the Fermi energy is usually found rather close to the CBM, these surface resonances at the CBM have an important implication because they can capture electrons from the CB in the bulk and act as acceptors as proposed by Klein [264]. As a consequence (111) and (011) surfaces of In_2O_3 should exhibit an electron depletion at high Fermi energy. The effect, however, is weak because it is caused by an increased density of states of CB sates at the surface, only. Especially, at more reducing conditions this depletion is likely to be dominated by a second effect.

In Section 12.3.2 we have shown that oxygen vacancies are more easily formed on (111) and (011) surfaces, than in the bulk [131, 274] (formation energies are always positive in the bulk). This is indicated by the existence of reduced surfaces at low oxygen chemical potential (Fig. 70). It was pointed out that the nature of these reduced (111) and (011) surfaces can be described by an accumulation of oxygen vacancies. Oxygen vacancies are, however, also created at higher oxygen chemical potential in large numbers. In a full thermodynamic equilibrium the oxygen vacancy concentration is always significantly higher in these (111) and (011) surfaces when compared with the bulk. When surface vacancies are created, more five- and fourfold In is produced as well. In addition, a positive surplus surface charge devlopes since oxygen vacancies are donors. The increased donor concentration towards the surface layer should effectively lead to an electron accumulation in this material which has already been measured [95]. Such an accumulation layer is sensitive not only to the bulk Fermi level position but also to the degree of equilibration of the sample as well as the environmental conditions under which it is measured, because it is a consequence of surface reduction.

The heavily reduced (111) and (011) (Fig. 70) surfaces consist of vacancy clusters which produce increasingly deeper and overlapping states at the surface. It is intersting that there is no evidence of a lone-pair localization but an increased number of metallic bonds.

Also for oxygen, surface resonances are observed. They can be found close to the VBM, at (111) and (011) surfaces. Fig. 74 (left) shows the oxygen 2*p* states as they move toward the band-gap from the VB when oxygen becomes increasingly under-coordinated. The different lines in the figure correspond to different oxygen species at

the surface and the highest lying O2p-states correspond to the oxygen ions labeled O-I in Fig. 67. Because the Fermi energy is usually high in In_2O_3, these states always remain occupied and cannot cause band-bending effects.

The (001) surface

In the case of (001) surfaces, there are three major regimes when going from reduced to oxidized surfaces. Fig. 74 shows on the right the total density of states for the cation terminated, stoichiometric and fully peroxide covered surfaces. In all cases gap-states appear, however, originating from different species.

For the cation terminated surfaces In-5s derived and half-filled metallic bands are found within the band-gap and the effective oxidation state of the cations is 1.5 according to Bader analysis. It is difficult to estimate the position of these states with respect to the band edges as they extend throughout the band-gap within LDA. Since these states are only partially filled and the Fermi energy is usually found at or even above the CBM, these gap-states should have an acceptor character causing an upward band bending for this surface.

In the case of ITO, tin behaves differently when compared with surface indium. The Sn derived surface states on this metallic surface are related to the Sn lone-pair (Sn^{2+}) and located below the In related surface states from which it can captures electrons. This explains the segregation tendency of Sn at metallic surfaces (Sect. 12.6.2). Thus Sn^{2+} leads to a saturation of the surface states dampening the electron depletion effect which would otherwise occur at cation terminated (001) surfaces.

On the stoichiometric (001) surface the (filled) gap-states originate from the O-2p of the topmost oxygen atoms similar to (111) and (011) surfaces. In this case, however, they are located higher above VBM and clearly split off from the host states. This is because of the lower coordination number of oxygen (2-fold) in comparison with (111) and (011) surfaces (3- and 4-fold). The dispersion of the states is low as expected for oxygen. The width of the gap-feature arises because oxygen are located at different elevation normal to the surface resulting in a variation of the electrostatic potential.

In the case of the peroxide surface, the (filled) gap-states originate from the anti-bonding $p - \pi^*$-orbitals of the dimers. The bonding $p - \sigma$-orbitals can be found as additional features at the bottom of the VB. Finally, due to the close proximity of the oxygen in the dimers, their corresponding O-2s states split significantly (~ 5 eV). A more detailed discussion on the electronic structure of peroxide defects in In_2O_3 can be found in Ref. [258].

We stress that in In_2O_3 or ITO the Fermi energy is usually found rather high in the band-gap. In particular it is high in comparison with the oxygen related surface states of stoichiometric and peroxide surfaces which are shown in Fig. 74 (right). The introduction of empty states with this energetic position (oxygen atoms or molecules) would lead to their immediate occupation so that they act as acceptors. The energy gain associated with this addition of oxygen and the capture of electrons from the CB of the material has been demonstrated already in Sect. 12.6.1 and resulted in additional

surface structures and surface segregation of dopants. Therefore, also the oxygen-rich (001) surfaces of In_2O_3/ITO should exhibit an electron depletion at the surface in contrast to the findings of King *et al.*[320].

The (001) surfaces exhibit a chemically rather complex behavior so that we have to consider also the possibility of an electron accumulation under certain conditions. In the case of stoichiometric (001) surfaces with low Fermi energy (no ionosorption of oxygen possible) the predominant defect mechanism at the surface could be oxygen vacancy creation similar to the (111) and (011) surfaces. In this case an electron accumulation could be possible. We believe, however, that for the majority of cases it is more likely to encounter an electron depletion, which is in marked contrast to (111) and (011) surfaces which should predominantly exhibit electron accumulation.

12.7.2 *Ionization potentials*

Besides the nature and presence of gap states, different surfaces can be distinguished by their ionization potential I_p. This is due to the changing dipole at the surface. Identification of surface terminations on textured or epitaxial films by their ionization potential is a reliable experimental method, because it is not sensitive to band bending effects. Apart form the surface orientation it is mainly the stoichiometry of the surfaces, which influences the I_p. As it was shown in section 12.3.2, each surface exhibits its own specific stoichiometry variations.

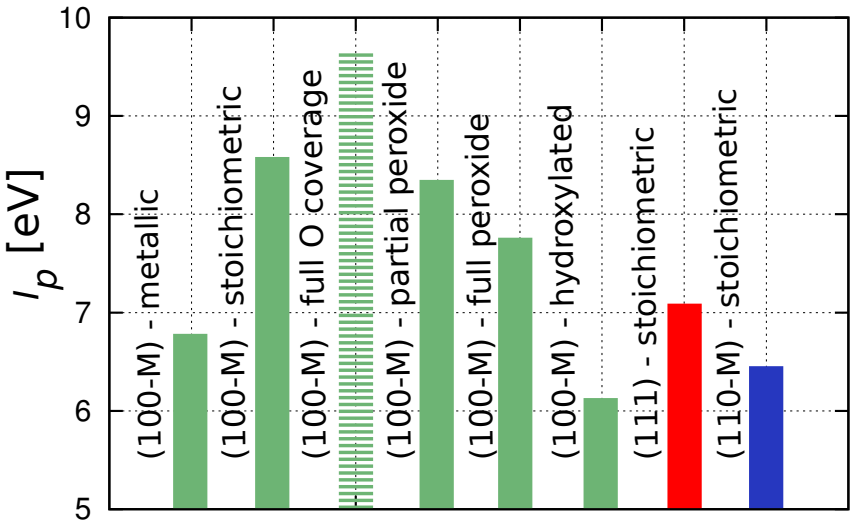


Figure 75: Calculated I_p for several In_2O_3 surfaces and different compositions.

Fig. 75 shows the calculated ionization potentials of different faces and surface compositions. The influence of the surface orientation and stoichiometry on the calculated I_p is as large as $\sim$ 3 eV. Among the stoichiometric surfaces the largest I_p is observed for the (001) orientation ($I_p \sim 8.5$ eV). The addition of oxygen increases the I_p significantly ($I_p \sim 9.5$ eV). Neither the stoichiometric nor the full oxygen terminated (001) surfaces are likely to be experimentally observable due to their high surface tensions. Surface reduction or oxygen dimerization stabilizes these surfaces and loweres their I_p. A hydroxylation leads to the lowest I_p ($\sim$ 6.1 eV). The values of Fig. 75 can be compared with those measured on epitaxial films [A10]. For films deposited at oxidizing conditions, the ionization potentials of 7.3 eV for (111) and 7.7 eV for (001) match well with the calculated values 7.7 eV for stoichiometric (111) and (001) full peroxide surfaces (7.1 eV), respectively. The excellent agreement of the values is remarkable and probably accidental to some extent. From the results of chapter 6, a significant underestimation of I_p can be infered ($\sim$ 0.5 eV) which worsens the agreement of the absolute values. The relative values, however, agee very well with the measurements. The results of this section imply that in polycrystaline films of In_2O_3/ITO large variations of I_p can be expected. The ionization potential partly determines the work function and, therefore, the injection barriers at contacts. Considering that electrodes for, *e.g.*, organic devices are often pretreated in oxygen-rich environments (ozone or oxygen plasma) [342], large inhomogeineties of injection properties should be assumed. A smaller variation could be achieved with reduced polycrystalline samples. In that case, however, the work function of the material would and be lower and would lead to high injection barriers in contact with organic materials [323].

12.7.3 *STM contrast*

The atomic arrangement on In_2O_3 surfaces has been studied by Morales *et al.* for the (111) [332] and (001) surfaces [329] using STM techniques. In the first case, they found a simple and stoichiometric bulk truncation in good agreement with their STM image simulation and our calculated phase diagram (Fig. 70). On the (001) surface they have found a considerably larger number surface features. The resulting contrast of the empty-states STM was interpreted on the basis of previous theoretical investigations [321, 337]. Specifically, what is termed as the (001-*M*) full peroxide termination in section 12.3.2 of the present study was first presented by Zhou *et al.* [337] and taken as the basic pattern for the interpretation of the STM images. Here, we discuss the essential STM features that can be expected for the various surfaces and specifically address the In_2O_3-(001) surface, for which experimental data is available.

Stoichiometric (011) surfaces

Generally, it is assumed that in empty states STM bright contrast mainly originates from the cations, since the CB is derived from extended In-5*s* states. However, the

actual atomic coordination of the surface cations is variable and has an additional effect on the relative brightness of the surface species.

The In-5*s* states tend to lower in energy when the coordination number of the cations is decreased. This is well illustrated in a simulated STM image of the stoichiometric (011) surface (Fig. 76) where fivefold and fourfold coordinated In can be found at essentially equal height in normal direction to the surface. For both, $(011\text{-}M)$ and $(011\text{-}D)$, the bright contrast originates only from fourfold In, whereas fivefold In is much darker. As expected, oxygen is invisible although it is located at similar altitude. Note that the Tersoff-Hamann (TH) STM image is obtained using a bias voltage of $V_{\text{bias}} = +1.8\ V$ and there is no great influence of the parameter on these essential features.

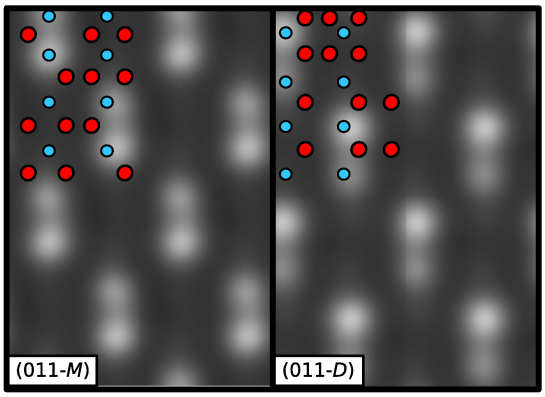


Figure 76: STM image simulation of stoichiometric $(011-M)$ and $(011-D)$ surfaces at +1.8 V. The atomic arrangement of the surface layer is shown.

(001) surfaces

In the case of the (001) orientation, additional difficulties in the interpretation of STM images arise because of the large variety of different surface phases (see Fig. 67), but also due to the alternating stacking sequence (see Fig. 66) with cations and anions at different heights. As a result there is a competition between surface topography and local electronic structure.

STM simulations for a collection of different (001) surfaces are presented in Figure 77 showing the diversity of observable patterns. The bias voltage was always kept at $+1.8\ V$. The positions of several atoms on the surface are marked in the images. Fig. 77 covers both, $(001\text{-}M)$ (first row) and $(001\text{-}D)$ (second row). For the last set of pictures (third row) in Fig. 77, Sn was present in the slab to illustrate the effect of *n*-doping.

It is well known that Sn dopants do hardly affect the electronic structure of In_2O_3 despite their large concentration. The dopants, however, drastically change the Fermi energy position of the material even to energies significantly beyond the CBM. This occupation of CB-states (Burstein-Moss shift) can lead to *unavailable* empty states for STM sampling and a potentially changed contrast. Generally, *n*-doping has a similar effect like an increased bias voltage since a larger fraction of higher lying states are sampled in both cases. This approximate analogy is also in line with the way the STM images change upon *n*-doping. Apart from the Fermi energy, Sn dopants have only a minor influence on the STM contrast and are not expected to be distinguishable from In.

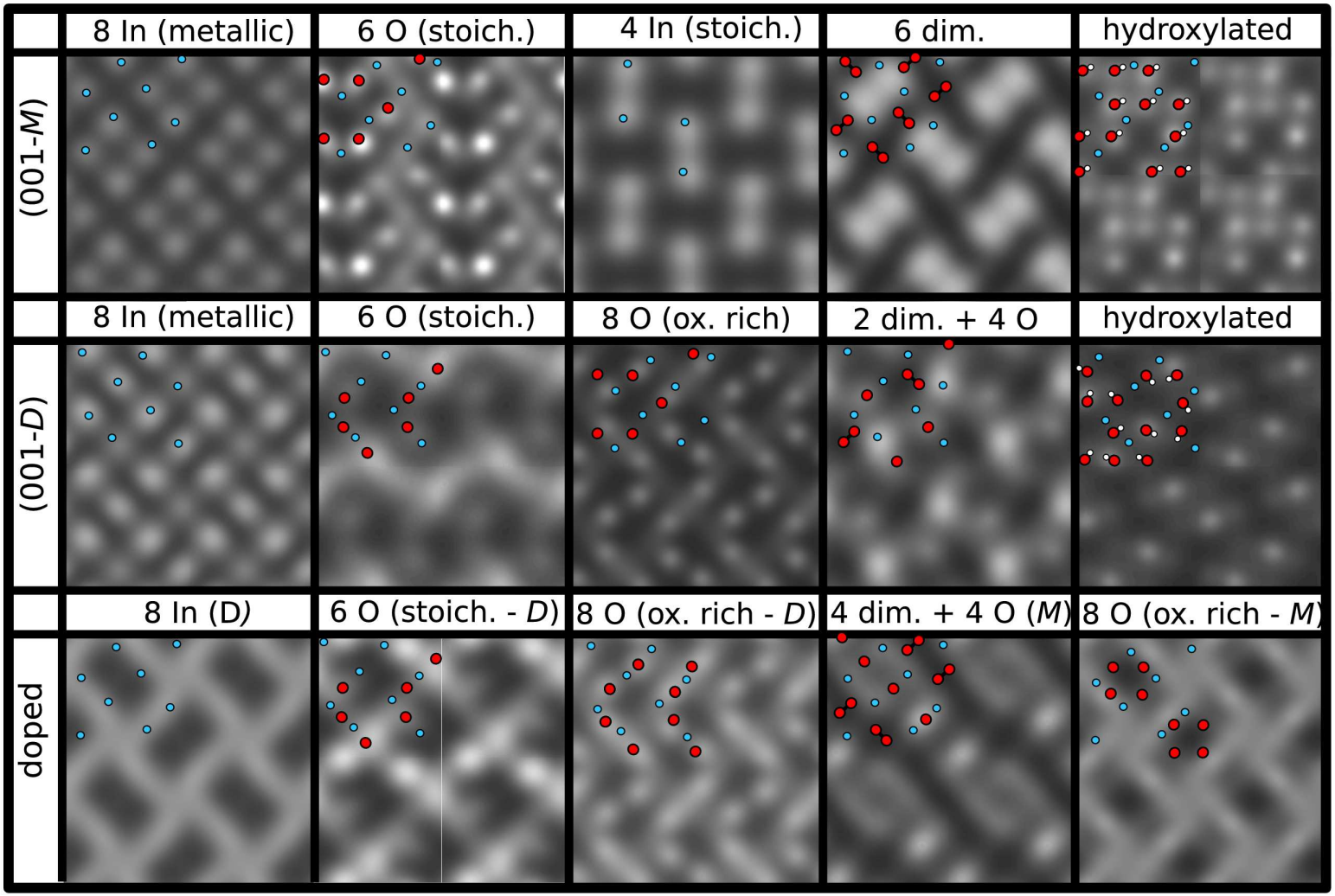


Figure 77: STM image simulation of various In_2O_3-(001) surfaces at a Bias of +1.8 V. The first, second and third rows contain $(001-M)$, $(001-D)$ and doped $(001-M/D)$ surfaces. The labels denote the content of surface species per cubic surface unit cell. The bulk truncations correspond to either 8 In or 12 O.

Metallic (001) surfaces

On the fully cation terminated surface, bright contrast simply indicate the location of the In cations. Already in this case, additional features appear between the In cations leading to slightly different appearance for $(001\text{-}M)$ and $(001\text{-}D)$, respectively. It is also noteworthy that these additional features are enhanced at higher voltages. Although the interpretation seems to be simple, the observed contrast is rather surprising because this surface has partially occupied In-5*s* states deep in the band gap. Since the Fermi energy is generally high in In_2O_3/ITO, these surface states may be fully occupied and lead to a very different contrast as it is depicted in the third row of Fig. 77. No single atoms can be identified but bright bands run along the $\langle 110 \rangle$ direction. This is the effect of the more directional In-5*p* contribution to the empty states.

Stoichiometric (001) surfaces

When oxygen is brought onto the terminal cation layer it is tempting to assume the height difference of > 1.2Å between oxygen and indium planes sufficient to produce a bright contrast by only the oxygen species located in the topmost layer. As a general result such an assumption cannot be sustained on the basis of the STM image simulations.

In the case of the stoichiometric terminations (6 oxygen per surface unit cell) some oxygen can be associated with bright features, but not exclusively. A large amount of the bright contrast originates from the underlying cation layer. The resulting images are generally difficult to interpret and change massively with the actual oxygen distribution on the terminal layer. This is important since for thermodynamical reasons a disorder is expected for theses surfaces. The calculated images show only the lowest energy structures which we could identify.

Common to many of these structures are, however, similar bright cross-like features as found for the *n*-doped metallic terminations. These features are often only rudimentarily present and appear as zigzag-bands running along $\langle 001 \rangle$. Again, these features tend to be enhanced at higher voltages but also by *n*-doping.

In summary, the competition of electronic structure versus the topography in the case of stoichiometric oxygen terminated surfaces is determined by the detailed arrangement of the surface oxygen as well as the sampling parameters making an unambiguous identification difficult. For comparison, the lowest energy half-indium termination is also shown in the first row of Fig. 77. The contrast produced by this surface suggest that it can very well be distinguished from the stoichiometric half-oxygen reconstructions.

Peroxide (001) surfaces

For undoped In_2O_3 and on the oxygen-rich side, the full or partial peroxide surface terminations are predicted to be thermodynamically favorable. In comparison to oxide anions on the stoichiometric surfaces, the peroxide anions undergo a more pronounced

inward relaxation bringing them closer to the underlying cation plane. As a result, the peroxide terminations exhibit no bright contrast at the location of the peroxo-ions anymore, whereas it was partially the case for the oxide anions on stoichiometric surfaces. Additionally, the splitting of the σ-states in the covalent bond of the dimers result in oxygen states that are inaccessibly high for tunneling. While the peroxide dimers constitute the top-most layer, the bright contrast originates form the indium cations located in the second layer. This behavior does not change with a varying bias or sampling height. In fact, the dimers are located within the darkest areas of the image. The STM images of surfaces containing both, peroxo anions and oxide anions, (2 dim + 4O and 4dim + 4 O) illustrate directly the relative brightness of the two anionic species. For example, the surface containing 4 dimers and 4 oxide anions is created by splitting two dimers from the full peroxide termination. Additional bright features appear at the locations where the dimers were split.

In the case of *n*-type doping, surfaces with less peroxide species but more oxide anions are thermodynamically more stable (see Sect. 12.6). Also for these non-charge balanced oxygen-rich surfaces, STM images are shown. Such surfaces are only stabilized by *n*-doping (8O-D/M, 4dim+4O). The images are characterized again by cross- and zigzag-like pattern and the contrast depends on the actual arrangement of the oxygen species. From thermodynamic considerations, these structures should be among the most stable surfaces for ITO under oxygen-rich conditions. Also in this case, however, it should be noted that the images are not unique due to the disordered nature of these terminations.

Hydroxylated (001) surfaces

In the case of the hydroxylated surfaces for both, $(001-M)$ and $(001-D)$, the interpretation is straightforward because the topmost hydrogen cations are bright. Since the hydrogen is tightly attached to oxygen of the underlying oxygen layer the hydrogen *brightens up* the oxygen ions which are located in their bulk-like arrangement. Nevertheless, $(001-M)$ and $(001-D)$ surfaces appear distinct which is due to the structure of the underlying cation layer. While on $(001-M)$ all hydroxyl-groups are at similar elevation, there are four protruding hydroxyls on $(001-D)$ surfaces. This difference results in twelve and four bright spots on $(001-M)$ and $(001-D)$, respectively.

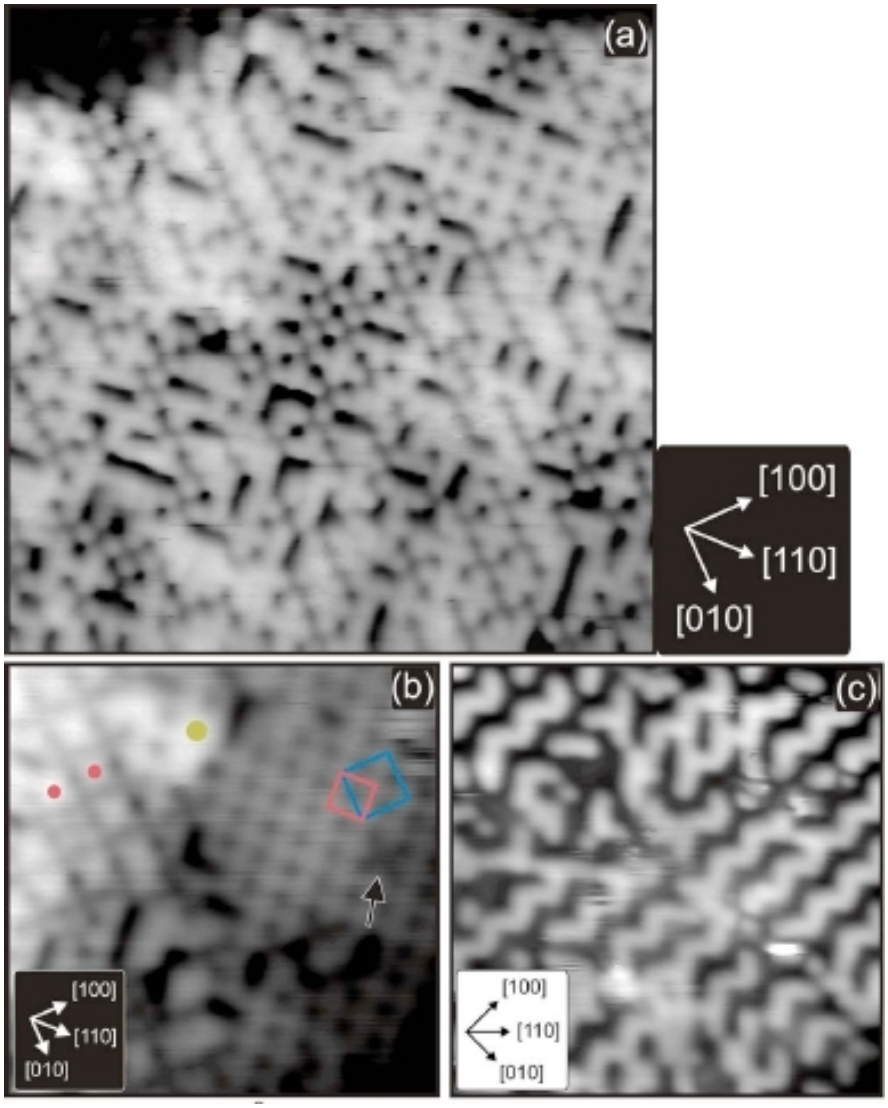


Figure 78: STM images of ITO-(001) films with different Sn content. (a), (b) 15.6 at% Sn, t=3.8 nm (a.) (15×15) nm^2, $I_{sample} = 1.1$ V, I_{tunnel}=0.78 nA; (b) 10×10 nm^2, $V_{sample} = +1.4$ V, I_{tunnel} =0.17 nA; (c) 3.6 at.% Sn, t=5.3 nm, 10×10 nm^2,$V_{sample} = +1.9$ V, $I_{tunnel} = 0.14$ nA.

Comparison with experiment

According to the STM image simulation, the fully hydroxylated (001), fully dimerized and half-indium termination and to lesser extent the full cation termination promise to have the highest degree of uniqueness for STM identification. These surfaces are not among those probably observed by Morales & Diebold [329] (reproduced in Fig. 78).

Their samples were exposed to oxygen-rich conditions prior to investigation and were highly doped with tin. Additionally, the contamination with water can be excluded. With a relatively high confidence the majority of observed features cannot be assigned to peroxide-type terminations as it was believed before.

From thermodynamic reasoning and with the aid of STM image simulations we are in the position to propose an alternative interpretation. In short, we suggest surfaces terminated mainly with oxide anion species and oxygen over-stoichiometry as described in Section 12.6. Under heavy *n*-doping these surfaces are thermodynamically most plausible. These surfaces exhibit also the largest energy gain by segregation of tin ions to the topmost cation layer. Heavily hydroxylated surfaces can be excluded on the basis of the preparation conditions but also the TH images. In addition, for heavily hydroxylated surfaces we have checked by calculation that there is no segregation tendency for Sn. In turn, the adsorption of hydrogen species on doped samples is less likely when compared with undoped In_2O_3 surfaces.

Besides the thermodynamics, the proposed surfaces also produce an STM contrast which is in qualitative agreement with many of the experimentally observed features (Fig. 78). Especially, the terminations labeled 8O-*M* and 8O-*D* closely resemble the observed structures with bright crosses and irregular zigzag bands. The surfaces found experimentally also involved a considerable degree of disorder which is only expected for this partially oxygen (oxide anions) terminated models. The occasional appearance of darker spots in the experiments could indicate the locations of remaining peroxide anions in the surface. Finally, we remark that the observed symmetry of the LEED pattern is also consistent with this interpretations assuming a high degree of disorder.

12.8 SUMMARY

We have investigated the thermodynamic stability of various surfaces and have found that the surface stability of stoichiometric In_2O_3 surfaces increases as $\gamma(111) < \gamma(011) < \gamma(211) < \gamma(001)$ which is consistent with observed shapes of nanoparticles [352] but not observed in the texturing of ITO films.

The most stable surfaces, ((111) and (011)) do not exhibit significant stoichiometric variations, whereas surfaces with polarity can be significantly off-stoichiometric. We observe the largest stoichiometry variations for the (001) surfaces which partially resemble the chemistry of the (211) surfaces and step edges of (111) surfaces. In_2O_3 can sustain oxygen-rich surfaces which owe their stability to peroxide formation. Under oxygen-poor conditions the surfaces exhibit metallic states often involving a strong overlap of In-5*s* states and in absence of a clear lone-pair formation at the

surface. While the most stable surfaces (111) and (011) are characterized by their chemical inertness, the (001) surfaces exhibit various surface phase transitions, which are sensitive to experimental boundary conditions.

Besides the chemical environment, the phase boundaries depend also on the in-plain strain. Because of the pronounced tensile surface stress due to the peroxide formation, (001) surfaces exhibit large energy variations as a function of an external strain that can originate from interfaces in thin film geometries. As a result, the phase boundaries are shifted, disfavoring the peroxide surface terminations and widening the range of the stoichiometric disordered phase.

Under the influence of hydrogen a different compensation mode for the polarity of the (001) surface becomes active. The fully hydroxylated and oxygen-rich surface can attain a surface tension even lower when compared with all other intrinsic In_2O_3-terminations. The resulting surface is obtained by one monolayer of dissociative water adsorption on the stoichiometric surface even at elevated temperatures. While a high degree of hydroxylation is expected for (001) surfaces, water adsorption on (111) but also (011) is weak and locally restricted to a small fraction of the total surface area. Nevertheless, dissociative water absorption is still possible also on those surfaces.

Alternative to hydrogen also other *n*-dopants influence the stability of the surfaces. In the case of extensive tin doping, new surface reconstructions appear as thermodynamic ground-states and are influenced by two effects. The direct effect of the dopants which are in contact with the surface species on the one side and the effect of the Fermi energy only, for which the exact location of the dopant is not relevant, on the other side.

The direct effect tends to destabilize the peroxide anions, whereas the higher Fermi energy generally stabilizes additional negatively charged species. Two new surface phases which should only be observable under *n*-doping have been presented in this context and more are believed to exist due to the continuous transition with increasing Fermi energy and the disordered nature of these reconstructions.

In light of the thermodynamic assessment of In_2O_3 and ITO surfaces, a new interpretation of available experimental STM data has been suggested. By explicit STM image simulation it was found that on (001) surfaces the bright contrast is provided mainly by indium cations, hydrogen and to only lesser extent by oxide anions but not by peroxide anions. These findings lead to the reinterpretation of the existing experimental STM images. Instead of the peroxide type reconstructions it appears more compelling to assume an off-stoichiometric (oxygen-rich) and disordered oxide ion termination, a surface which is not particularly stable in the case of pure In_2O_3 and only stabilized by *n*-doping.

As the most general and practical result the appearance of preferential (001) texturing of ITO samples was explained despite the highest surface tension of (001) surfaces in the case of pure In_2O_3. The combination of a high Fermi energy and oxygen-rich conditions together with the electrostatic energy gain due to Sn segregation and the probably high growth rates on (001) surfaces lead to the predominant (001) texturing of ITO samples.

12.9 CONCLUSION

In summary, we have extensively studied most of the experimentally observed In_2O_3/ITO surfaces and discussed the effects of environment (oxygen partial pressure and temperature), substrate mismatch as well as hydrogen and doping. Our results provide a comprehensive basis for future studies of In_2O_3 surfaces. Again, we emphasize the role of the Fermi energy on the surface stability, a parameter which is often neglected in related studies. Especially, when the variability of this parameter can be large as in In_2O_3 it can have a determining role on the observed surface features.

With this comprehensive study, the possibilities as catalyst, gas-sensing and electronic contact material can proceed in a more systematic way based on a sound microscopic approach. Especially, with the use of epitaxial In_2O_3/ITO, the rather distinct properties of different In_2O_3 surfaces can be exploited in the future.

For example, (001) surfaces provide a probably too large variability of surface composition in order to serve well as gas sensors with predictable gas-response in a wide range of conditions. On the other hand this variability is favorable for In_2O_3 being a combustion catalysis because the (001) surfaces can contain a large content of weakly bound oxygen species. In the context of contact materials not only the orientation [340] but also the surface composition can provide some variability for the ionization potentials and work functions [324].

Also in the field of nanoparticle research our thermodynamic assessment of the surfaces can be of help. For example, the predominant equilibrium nanoparticle shapes are octahedra due to the low (111) surface tensions. In order to produce more round-shaped nanoparticles it is advicable to obtain them at very reducing conditions where the surface energetics converge to very similar values for all surfaces (Fig. 70). In contrast, for cubic nanoparticles it is advisable to use n-doped In_2O_3 and in conjunction with either oxygen plasma or ozone exposure. Alternatively, synthesis routes from the hydroxide could also be successful in producing cubic particles with hydroxylated (001) faces.

SURFACES: SnO_2 VS. In_2O_3

We are now in the position to compare the surface properties of In_2O_3 and SnO_2. In the following we summarize the key findings and list important similarities and differences between the two materials.

The major similarities are:

- Stoichiometry variations at the surface are possible for both materials.
- Positive surface charging due to surface oxygen vacancy formation at mildly reducing conditions is found for In_2O_3-(111)/(011) and SnO_2-(110) surfaces.
- Surfaces show a tendency towards reduction in accordance with their bulk defect equilibria.
- Disordered surface reconstructions exist for In_2O_3-(001) and SnO_2-(110) surfaces.
- Electronic structure and STM contrast is similar for stoichiometric surfaces.

More interesting and predominant are, however, the differences:

- Generally, surface tensions of bulk truncated SnO_2 surfaces are slightly larger compared to In_2O_3 surfaces. A similar trend was found also for bulk defect formation energies (chapter 8).
- Surfaces show a chemically uniform behavior for SnO_2, whereas the chemistry on In_2O_3 surfaces depends on orientation. This trend also holds for the ionization potentials.
- In_2O_3-(111) and (011) surfaces are chemically inert, whereas all SnO_2 surfaces are reducible at moderately reducing environments.
- Only In_2O_3-(001) surfaces can be oxidized. In contrast, SnO_2 surfaces exhibit only oxygen deficiency. This trend is in accordance with the bulk defect equilibria of the respective materials.
- SnO_2 exhibits a lone-pair effect under reduction (under-coordination), whereas under-coordinated In tends to form metallic bonds with other In.

Part VI

CONCLUSION

13

SUMMARY

The defect thermodynamics and kinetics as well as the thermodynamics of surfaces has been studied in depth by DFT calculations. The key findings are summarized below.

DEFECTS THERMODYNAMICS

- The intrinsic defect thermodynamics is largely governed by the formation of oxygen vacancies in the reducing regime, whereas it is dominated by neutral peroxide defects in the oxidizing regime for both materials.
- SnO_2 is generally less affected by the presence of point defects. More generally, SnO_2 contains the lowest amount of intrinsic defects among the TCOs ZnO, In_2O_3, Ga_2O_3 and CdO.
- The ionization energy of the oxygen vacancies is greater in SnO_2 as compared to In_2O_3. As a result, the intrinsic carrier concentrations are lower in SnO_2.
- The formation of negative oxygen interstitials is only possible in In_2O_3 and at *n*-type doping conditions.
- The principal acceptor defect in In_2O_3 is the negative oxygen interstitial, whereas it is the tin vacancy in SnO_2.
- The presence of cation interstitials and associates comprising cation interstitials can be excluded on the grounds of high formation energies for SnO_2 and In_2O_3.
- Cation interstitials in SnO_2 occur in a split interstitial configuration instead of occupying regular interstitial sites.
- The formation of double oxygen vacancies at reducing conditions is important only for SnO_2.
- Defect formation entropies can be large in TCOs. They can change the order of stability among point defects. Indium vacancies, for example, are favored over oxygen interstitials as acceptor defect at high temperatures in In_2O_3.

- Defect formation entropies at constant pressure conditions are largely goverened by the strain fields around the defect.
- Defect formation entropies in TCOs are generally larger for acceptor than for donor defects.
- Hybrid functionals remedy the band gap error encountered in LDA or GGA calculations.
- Correction schemes previously applied to the band gap were in error for two distinct reasons: First, a significant band gap correction was indicated for the VB states by the hybrid functional calculations, while previous corrections acted mainly on the CB. Second, the defect relaxation energies are significantly affected by the exchange-correlation functional, an energy contribution which cannot be corrected by band edge shifting. Hybrid functionals increase the relaxation energy of V_O^{2+}.
- Hybrid functionals predict the charge transition levels of oxygen vacancies by up to 1.5 eV higher in the band gap as compared to LDA/GGA values.
- In the case of In_2O_3 the hybrid functionals indicate the charge transition energy of oxygen vacancies in close proximity to the CB, making thermal ionization likely. In SnO_2 and especially ZnO the transition levels are located deeper within the band gap but significantly higher than previously suggested.
- The band gap underestimation is only marginally caused by the so-called $p-d$ repulsion in In_2O_3 and even less in SnO_2. As a result, thermodynamic calculations are valid for In_2O_3 and SnO_2, even without explicitly considering the effect of d-electrons.
- In the hybrid functional description, formation energies of donors are generally lower (up to 6 eV). For all oxides studied, large and negative formation energies of donor defects in the lower part of the band gap were identified even under oxidizing conditions. A technologically important implication of the lower donor formation energies compared to previous reports is that no equilibrium conditions exist at which p-type conductivity can be sustained for In_2O_3, SnO_2 and ZnO.
- On the basis of first-principles calculations, the maximum free electron concentrations are higher in SnO_2 compared to In_2O_3. This is at odds with experimental trends. As we made sure that the computational setup is not likely to alter this conclusion, the doping limitations in SnO_2 in turn can be accounted to the specific dopants presently used. We conclude that alternative impurities could exceed the levels presently achieved by the conventional Sb or F dopants.

DEFECTS KINETICS

- The double positive oxygen vacancies have migration energies in the expected range ($\sim$ 1 eV) for both materials.
- In undoped and acceptor doped material V_O^{2+} mediates oxygen diffusion in SnO_2 ($E^M = 1.3$ eV) and In_2O_3 ($E^M = 1.0$ eV). This finding and the calculated migration energies are in good agreement with experiments.
- Oxygen diffusion is mediated by neutral peroxide oxygen dumbbell interstitials at the highest oxygen chemical potentials for SnO_2 ($E^M \sim 0.7$ eV) and In_2O_3 ($E^M = 1.2$ eV). In contrast to V_O^{2+} this defect cannot be driven by electric fields.
- Only in the case of *n*-doped In_2O_3 diffusion can be mediated by negative oxygen interstitials.
- In In_2O_3 the migration energy of negative oxygen interstitials is larger ($E^M = 1.4$ eV) when compared with V_O^{2+} despite the relatively open crystal structure. As a result, oxygen is relatively stable within ITO samples and cannot diffuse out easily from contacts in, *e.g.*, organic electronic devices.
- Cation vacancies migrate more slowly when compared with oxygen vacancies and interstitials in In_2O_3 and SnO_2. In light of the smaller cation radii compared to anions this is surprising and contrasts with the behavior of other TCOs like ZnO, CdO and Ga_2O_3.
- In SnO_2 cation interstitials do not migrate preferentially within the interstitial channels present in the rutile structure in $\langle 001 \rangle$ direction. Instead, they form a split-interstitial configuration and move within the (001) plane.
- Low and almost vanishing energy barriers were identified for cation interstitials in both, In_2O_3 and especially SnO_2. In the case of In_2O_3 these low energy transition states do not lead to an overal fast diffusion due to geometric aspects of the underlying crystal structure. In the case of SnO_2 fast diffusion is strongly anisotropic and only possible perpendicular to the *c*-axis.
- For diffusion within the C-type rare earth structure (bixbyite) of In_2O_3, the lowest energy transitions are not a fingerprint of the effective migration parameters like for structures with smaller unit cells (*e.g.* SnO_2). Despite low barriers the defects cannot escape from certain locations within the unit cell without transitions over high energy barriers. This general finding has similarly important implications for other oxides of this structure type like Y_2O_3, Er_2O_3, Tm_2O_3... etc.
- Light irradiation with energies close to the band gap excitation can lead to an enhancement of oxygen diffusion in In_2O_3 and especially in ITO. This effect originates from the structural similarity between the saddle point configuration

of O_i to $O_{i,db}$ defects. Similar effects are not possible in the case of SnO_2 due to the low abundance of O_i.

SURFACES

Finally, four different surfaces of both materials were studied in detail. The most interesting behavior was observed for the (110) faces of SnO_2 and the (001) faces of In_2O_3.

- Surface tension of SnO_2 surfaces tend to be higher compared to In_2O_3 surfaces.
- The variability of surface composition increases for surfaces with a higher surface tension.
- Reduced SnO_2 surfaces are generally characterized by a layer of SnO-like phase at the surface. This is usually achieved by removing the topmost oxygen layer.
- The SnO_2-(110) surface shows a more complex behavior: In this case Sn^{2+} surface interstitials are the essential motif for understanding the complex reconstructions at reducing conditions. This is remarkable since tin interstitials do not play an important role for the bulk defect equilibria of SnO_2.
- Mild reduction of SnO_2-(110) surfaces results in creation of oxygen surface vacancies and Sn_2O_2 surface defects comprising tin surface interstitials.
- At lower oxygen chemical potential Sn_2O_2 join together and form strands in various and disordered arrangements. This explains the (1×1) high temperature reconstruction in excellent agreement with experimental observations.
- The second high temperature reconstruction is obtained by adding a second building block (Sn_4O_6) to the Sn_2O_2 units. The mixing of the two motifs is possible in a disordered manner and results in a second disordered reconstruction with an effective (1×2) periodicity.
- By testing several previously proposed models of metastable (4×1) reconstructions we could show that none of them is consistent with experimental findings. In contrast, we have proposed a candidate for the meta stable $c(2 \times 2)$ reconstruction.
- In contrast to SnO_2, In_2O_3 surfaces exhibit a chemically more complex and non-uniform behavior. Surface oxidation and reduction is possible.
- (011) and (111) surfaces of In_2O_3 are chemically inert up to low oxygen chemical potentials close to the materials stability limit.
- In contrast, In_2O_3-001 surfaces show a large variability of surface composition.

- STM images for ITO-(001) surfaces cannot be explained by considering the ground-state surface reconstructions of In_2O_3. Instead, the effect of the dopant has to be included.
- At high Sn doping levels, oxygen-rich and non-peroxide surfaces are stabilized. The crucial factor is the increased Fermi level in the system. These reconstructions explain many important features observed in STM.
- Without doping, the (001) surfaces have the lowest stability among all In_2O_3 surfaces. The increased Fermi level and dopant segregation are effective in reducing the surface tension of (001) surfaces. Segregation of Sn to other surfaces is less likely. These mechanisms are the origin of the predominance of (001) oriented surfaces in ITO films.
- Alternative to *n*-doping, hydroxylation of the polar (001) surface is effective in reducing the surface energy to values comparable or even lower than other faces. This is because water adsorption is relatively weak on the more stable (011) and (111) surfaces.
- In_2O_3-(111) and (011) surfaces exhibit electron accumulation layers as a consequence of surface reduction.
- In_2O_3-(001) surfaces show a porpensity towards electron depletion layers.

14

OUTLOOK

This final chapter deals with the specific implications which emerge from the results of this thesis.

INTRINSIC CONDUCTIVITY

The origin of the intrinsic conductivity in TCOs is still an ongoing question and should further be investigated. The results of the present thesis provide evidence of lower formation and ionization energies of oxygen vacancies than previously anticipated. The presence of hydrogen in large amounts was postulated for all TCOs as an alternative electron donor [242, 353, 132, 134, 354]. Using IR absorption [253] and muonium spin rotation and relaxation techniques [355, 356], there is also experimental confirmation. As a general conclusion the most difficult case is encountered in TCOs where both, hydrogen and oxygen vacancies can contribute significantly to the conductivity depending on environment. The presently most pressing question is, therefore, to quantitatively identify their relative contributions for a given thermodynamic environment. This task is further complicated by the reducing effect of hydrogen and the possibly metastable character of the oxygen vacancies. Experimentally, a key to this understanding could be the differences in defect kinetics for the two defect species. Compared to the calculated migration energies of hydrogen interstitials ($\sim$ 0.6 eV [132]), the vacancies in TCOs are slow ($E^M > 1$ eV) (chapters 9 and 10). A combination of kinetic experiments at high and low temperatures could successfully divide the relative contributions of the two species. In addition, further refinements of the thermodynamic parameters of the vacancies are necessary by using even more advanced electronic structure methods such as many-body perturbation theory.

A second question, which will need further clarification, is the absolute magnitude of measured conductivities and non-stoichiometry in comparison with theoretical predictions. Especially, for samples at higher oxygen chemical potential, the formation energies of vacancies cannot explain the high values of conductivity and oxygen deficiency alone. A key to the understanding of this behavior could be the properties of surfaces and interfaces. As we have shown in chapters 11 and 12, surfaces of SnO_2 and In_2O_3 can be reduced more easily when compared with the bulk. For some surfaces this reduction is associated with an electron accumulation layer. In

materials with small grains or in thin films, the conductivities could therefore be overestimated. For example, a thermogravimetric investigation of SnO_2 [123] showed that the loss of oxygen from the sample is largely connected to the surface, whereas the bulk non-stoichiometry remained lower but measurable and in better agreement with predicted values. Similar contributions could arise from grain boundaries (GB). The typical grain sizes of the materials employed are often on the nanoscale so that the consideration of such effects is essential. The properties of GBs can be assessed by electronic structure methods in a very similar manner like for surfaces. Certain high symmetry configurations of GBs can be easily constructed in periodic boundary conditions and their interface energies and equilibrium stoichiometries can be obtained using the formalism of section 4.6.

EXTRINSIC DOPING

Generally, the performance of extrinsically doped TCOs is impressive. ITO is the most frequently used TCO in many technologies and has reached the bounds of its performance. Unfortunately, a high market volatility and risk is connected with ITO. In addition, one material alone will not be able to cover the demands for all applications. For example, a high transparency and conductivity may be needed in conjunction with a good chemical resistivity, or simply a specific crystal structure in order to provide epitaxial growth. Therefore, there is a general demand for diversification of TCO materials. In this context, the present study has made a considerable contribution as it could be shown by a detailed theoretical analysis that for SnO_2 based materials a further increase of the electric conductivities is possible. SnO_2 is a chemical and wear resistant material which, at low temperatures, is able to block oxygen exchange at its surface [291]. These properties add functionalities to devices which are not given by ITO. In order to reach the levels of conductivity in SnO_2 comparable to that of ITO, new dopants and dopant combinations have to be explored such as the transition metals Ta, Nb, W, Mo, V. First attempts by groups in Korea [357] show promising results. The suitability of dopants in terms of ionization, their kinetics or solubility and segregation tendency to surfaces/interfaces can be mapped out using electronic structure methods prior to material production.

CONTACT PROPERTIES

The diverse surface structures found in the case of In_2O_3 raise the question, in how far this material is a suitable contact material. Based on the results of chapter 12 a large lateral variation of injection barriers has to be expected for polycrystalline ITO. This is because of the distinct properties of different In_2O_3 surfaces as a function of the oxygen chemical potential. In contrast, SnO_2 surfaces exhibit a uniform behavior and different surfaces undergo very similar phase transitions (chapter 11). Furthermore, no oxygen-rich surfaces are possible for SnO_2. Polycrystalline films based on SnO_2

are, therefore, likely to exhibit more uniform injection properties, as well. The major drawback of SnO_2 is still the lower conductivity. According to our results this is an issue which could be resolved in the future. As far as the suitability for electronic contacts is concerned, SnO_2 appears superior compared to ITO once the conductivities are improved. Besides more homogeneous injection properties doped SnO_2 is not expected to contain excess oxygen in the bulk and oxygen diffusion is generally slower (chapters 7 and 10).

GAS SENSING AND CATALYSIS

The present thesis provides a detailed understanding of the structure and chemistry of the surfaces in In_2O_3 and SnO_2. This microscopic knowledge is the basis for the investigation of gas response and catalytic activity at the surfaces of these materials. A feature common to all previous surface studies on TCOs is that surface adsorbates or other surface species were treated only in the charge neutral case. Considering the literature and the results of this thesis, there is strong evidence that in most cases this assumption is not justified. Especially, the gas response effect relies explicitly on charging effects. In such cases the adsorption and formation energies of surface species carry a Fermi energy dependence similar to charged point defects. For a consistent description, the grand canonical formalism of electrons can be used in the same way as it was done for defects in this thesis. The formation (adsorption) energies of various surface species can be calculated in different charge states and plotted as a function of the Fermi energy. This can be done not only at the surface but also in several subsurface layers until the numbers converge to the bulk values. When combined with barrier calculations this information can be used in conjunction with differential equations or kinetic Monte-Carlo methods to predict the space charges at various surfaces as a function of environment and enables a first-principles modeling of the gas response and band bending. In the same way, valuable information can be obtained for the field of catalysis.

METASTABLE PHASES

A distinct feature of SnO_2 in comparison with other TCOs is the existence of a stable suboxide. Though experimentally easily accessible, this compound is not necessarily an equilibrium phase in the Sn-O phase diagram [72] similar to the mixed valence compounds Sn_3O_4 and Sn_2O_3.

Fingerprints for the formation of suboxides on smaller scale can be found in both, the defect equilibria and surface structures. In the bulk, this is manifested in the tendency to form double oxygen vacancies (chapter 7). At surfaces the formation of Sn^{2+} species is common for all terminations at reducing conditions. Because of the close connection of these structures with SnO or the intermediate oxides, the existence

of other oxygen deficient (meta) stable structures can be suspected. In the bulk such structures could even have a higher probability of survival due to slower kinetics.

A prime example of such metastabilities is given by the (4×1) reconstruction of SnO_2-(110) surfaces. This structure has been observed by several groups and is the most reproducible surface reconstruction of SnO_2-(110) faces. Its structure is yet unresolved but remarkably stable up to temperatures as high as $\sim 900°C$. In the same spirit Seko *et al.*[72] proposed the existence of a homologous series of intermediate tin oxides. The pattern discovered by Seko *et al.* can be generalized to construct planar defects or twins in (101) planes*. Further investigations of such oxygen deficient structures could be the key to understand not only the SnO_2-(110) (4×1) reconstruction but also differences to other TCOs like In_2O_3 and ZnO where suboxide formation is less likely. Moreover, it is plausible that the formation of these structures is not only triggered by reducing environments or non-equilibrium, but also by high Fermi energies. As a result, the doping difficulties in SnO_2 could be related to defects of this kind. Preliminary results obtained on this topic show promising results. Similar ideas can also be tested for In_2O_3 and other TCOs.

*Similar constructions are also possible in other planes.

ERKLÄRUNG – DISCLAIMER

Die vorliegende Arbeit wurde im Zeitraum von Juni 2007 bis Mai 2011 im Fachgebiet Materialmodellierung am Institut für Materialwissenschaft der Technischen Universität Darmstadt bei Herrn Prof. Dr. rer. nat. Karsten Albe angefertigt.

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Hilfsmittel angefertigt habe. Von mir wurde weder an der Technischen Universität Darmstadt noch an einer anderen Hochschule ein Promotionsversuch unternommen.

Darmstadt, den 18. Mai 2011

Péter Ágoston

DANKSAGUNG – ACKNOWLEDGMENTS

Ohne die Hilfe einer ganzen Reihe von Menschen wäre mir die Erstellung der vorliegenden Arbeit nicht möglich gewesen.

Insbesondere danke ich Professor **Karsten Albe** für das Thema und den Titel dieser Arbeit, sein Vertrauen, stetige Unterstützung sowie enorme Freiheiten in Sachen Arbeitszeiten und Forschungsthemen.

Ich danke Professor **Heinz von Seggern** für das koreferieren dieser hoffentlich nicht zu langen Arbeit, sowie **Wolfram Jägermann** und Prof. **Rolf Schäfer** für das Lesen und die noch anstehende Fragestunde.

Besonderen Dank schulde ich Prof. **Andreas Klein** für reichliche wissenschaftliche Diskussionen, wertvolle Hinweise und gute wissenschaftliche Kooperation.

I would like to thank Prof. **Risto Nieminen** and Prof. **Martti Puska** for warmthly hosting me at Alto University as well as their advice and cooperation in joint research projects.

Besonders herzlich möchte ich mich bei meinen **Eltern**, meiner **Schwester** und **Janina** für ihre stetige Unterstützung bedanken.

Ich danke den **M & M's** für die entspannte und oft serienreife Arbeitsatmosphäre, insbesondere **Renate H.**, die mich vor dem Chaos bewahrt hat, **Yvonne R.** für Motivation und Korrekturenlesungen, **Johann P.** und **Manuel D.** für interessante Diskussionen über Thermodynamik und viele andere Themen, **Melanie G.** für chemikalische Zusammenhänge sowie **Daniel S.** für Hilfe bei Kummer den man nur im Osten versteht.

Ich bedanke mich für die erfolgreiche Zusammenarbeit und für erleuchtende Diskussionen bei **André W.**, **Christoph K.** und **Mareike H.** aus dem Fachgebiet Oberflächenforschung.

Many thanks to **Javad H.** for a great and productive time in Helsinki.

Nicht zuletzt möchte ich mich bei all den Leuten bedanken die für das Erfolgreiche Zustandekommen dieser Arbeit beigetragen haben aber namentlich nicht erwähnt

wurden.

This work was funded by the *Sonderforschungsbereich 595* "Fatigue of functional materials" of the "Deutsche Forschungsgemeinschaft". The project was partially supported through a bilateral travel program funded by the German foreign exchange service (DAAD). Grants of computing time at HHLR supercomputers at HRZ, FZ Juelich and CSC Espoo are kindly acknowledged.

Curiculum Vitae

Personal data

Name	Peter Agoston
Address	Mailänder Strasse 12 Appartement 2004 60598 Frankfurt am Main
Tel.:	+491793202661
E-Mail	agoston@mm.tu-darmstadt.de
Birth	31. October 1980 Tirgu Mures (Romania)
Sex	male
Age	30
Family status	single, no children

University Education

07/2011	Defense of PhD-Thesis
since 07/2007	**PhD Thesis:** *"A Comparative study of point defects and surfaces in In_2O_3 and SnO_2 by first-principles methods"* Materials Modeling Division (Prof. K. Albe), Materials Science, Darmstadt University of Technology, Germany
11/2006–05/2007	**German Diploma Thesis:** *"Electronic Structure and Thermodynamics of Intrinsic Point Defects in Indium Oxide"* Materials Modeling Division (Prof. K. Albe), Materials Science, Darmstadt University of Technology, Germany
09/2001–10/2006	**Studies:** Materials Science, Darmstadt University of Technology, Germany

Professional Experience

07/2007–Present **Research Associate (Wissenschaftlicher Mitarbeiter)**:
Quantum mechanical modeling of oxide materials (PhD), teaching, supervision of diploma/master students and research assistants
Materials Modeling Division (Prof. K. Albe), Materials Science, Darmstadt University of Technology, Germany

11/2006–05/2007 **Research Assistant:**
Quantum mechanical modeling of oxide materials (Diploma Thesis)
Materials Modeling Division (Prof. K. Albe), Materials Science, Darmstadt University of Technology, Germany

Professional Experience Abroad

11/2008–03/2009 **Visiting Scientist:**
Quantum mechanical modeling of solids,
Department for Applied Physics and Technology (Prof. Risto Nieminen)
Alto University, Espoo, Finland

04/2005–10/2005 **Internship:**
Investigation of plastic deformation in alumina using electron beam techniques (EBSD, TEM)
Institute of Material Science (Prof. M. Hoffman),
University of New South Wales, Sydney, Australia

Civil Service

09/2000–06/2001 Kinderheim Reinhardshof (Frankfurt am Main)

School Education

07/1982–07/2000 Secondary School: Georg Büchner Gymnasium, Bad Vilbel, Germany

07/2000 Graduation from secondary school (*Abitur*) Georg Büchner Gymnasium, Bad Vilbel, Germany

Scholarships

10/2004 Member of the German National Merit Foundation

2005 Dr. Jürgen Ulderup Scholarship

Oral Presentations at International Conferences

P. Ágoston, A. Klein, K. Albe
Thermodynamics and Kinetics of Intrinsic Point Defects in In_2O_3
International Conference on TCO Materials, Heraklion, Crete, 2008

P. Ágoston, P. Erhart, A. Klein, K. Albe
Ab-initio modelling of thermodynamics and kinetics of point defects in indium oxide
DPG Spring Meeting, Berlin, 2008

P. Ágoston, A. Klein, K. Albe
First-principles calculations on self-diffusion in indium oxide
DPG Spring Meeting, Dresden, 2009

P. Ágoston, P. Erhart, A. Klein. R.M. Nieminen, M.J. Puska, K. ALbe
Comparative study of diffusion in n-type TCO materials
DPG Spring Meeting, Regensburg, 2010

P. Ágoston, K. Albe
Intrinsic n-type behavior in transparent conductive oxides: A comparative hybrid-functional study of In_2O_3, SnO_2, and ZnO
DPG Spring Meeting, Regensburg, 2010

P. Ágoston, A. Klein, R.M. Nieminen, M.J. Puska, K. Albe
Intrinsic Defects of Transparent Conducting Oxides: A Comparative Hybrid-Functional Study of In_2O_3, SnO_2 and ZnO
5th forum on New Materials, Montecatini Terme, Italy 2010

P. Ágoston, C. Körber, A. Wachau, A. Klein, R.M. Nieminen, M.J. Puska, K. Albe
Doping mechanisms in SnO_2: Insights from first principles calculations
MRS Fall Meeting Boston 2010

Book Chapters

A1 K. Albe, P. Ágoston and J. Pohl, in *Advanced Characterization Techniques for Thin Film Solar Cells*, edited by D. Abou-Ras, T. Kirchartz and U. Rau (Wiley-VCH, Berlin, 2011), Chap. Ab initio modeling of semiconductors.

Peer Reviewed Articles

A2 P. Ágoston and K. Albe, *Formation entropies of intrinsic point defects in cubic In_2O_3 from first-principles density functional theory calculations*, Phys. Chem. Chem. Phys. **11**, 3226 (2009).

A3 C. Körber, P. Ágoston and A. Klein, *Surface and bulk properties of sputter deposited undoped and Sb-doped SnO_2 thin films*, Sensors and Actuators B: Chemical **139**, 665 (2009).

A4 J. Suffner, P. Ágoston, J. Kling and H. Hahn, *Chemical vapor synthesis of fluorine-doped SnO_2 (FTO) nanoparticles*, J. Nanopart. Res. **6**, 1345 (2009).

A5 P. Ágoston, P. Erhart, A. Klein and K. Albe, *Geometry, electronic structure and thermodynamics of intrinsic point defects in indium oxide*, J. Phys. Condens. Matter **21**, 455801 (2009).

A6 P. Ágoston, K. Albe, R. M. Nieminen and M. J. Puska, *Intrinsic n-Type Behavior in Transparent Conducting Oxides: A Comparative Hybrid-Functional Study of In_2O_3, SnO_2, and ZnO*, Phys. Rev. Lett. **103**, 245501 (2009).

A7 P. Ágoston and K. Albe, *Ab initio modeling of diffusion in indium oxide*, Phys. Rev. B **81**, 195205 (2010).

A8 P. Ágoston, C. Körber, M. J. Puska, R. M. Nieminen, A. Klein and K. Albe, *Limits for n-type doping in In_2O_3 and SnO_2: A theoretical approach by first principles calculations using hybrid-functional methodology*, J. Appl. Phys. **108**, 053511 (2010).

A9 P. Ágoston and K. Albe, *Disordered reconstructions of the reduced SnO_2-(110) surface*, Surf. Sci. **605**, 714 (2011).

A10 M. V. Hohman, P. Ágoston, A. Wachau, T. Bayer, J. Brötz, K. Albe and A. Klein, *Orientation Dependent Ionization Potential of In_2O_3: A Natural Source for Inhomogeneous Barrier Formation at Electrode Interfaces in Organic Electronics*, J. Phys.: Condens. Matter **In Press**, (2011).

A11 C. Korber, A. Wachau, P. Ágoston, K. Albe and A. Klein, *Self-limited oxygen exchange kinetics at SnO_2 surfaces*, Phys. Chem. Chem. Phys. (2011).

A12 P. Ágoston, K. Albe, R. M. Nieminen and M. J. Puska, *Reply to the Comment by S. Lany and A. Zunger*, Phys. Rev. Lett. **106**, 069602 (2011).

A13 P. Ágoston and K. Albe, *Thermodynamic stability, stoichiometry and electronic structure of bcc-In_2O_3 surfaces*, Phys. Rev. B (2011), accepted.

A14 P. Ágoston and K. Albe, *Thermodynamics and kinetics of intrinsic point defects in SnO_2*, Solid State Ionics (2011), submitted.

Frankfurt am Main, 11. October 2011

BIBLIOGRAPHY

[1] *Fatigue of functional materials* (DFG, TU-Darmstadt, 2010).

[2] J. C. Scott, J. H. Kaufman, P. J. Brock, R. DiPietro, J. Salem, and J. Goitia, J. Appl. Phys. **79**, 2745 (1996).

[3] D. S. Ginley, H. Hosono, and D. Paine, *Handbook of Transparent Conductors* (Springer, New York Heidelberg Dordrecht London, 2010).

[4] G. Thomas, Nature **389**, 907 (1997).

[5] K. Chopra, Thin Solid Films **102**, 1 (1983).

[6] R. Gordon, MRS Bull. **25**, (2000).

[7] C. G. Granqvist, Sol. Energ. Mat. and Sol. C. **91**, 1529 (2007).

[8] Y. Sawada and Y. Taga, Thin Solid Films **116**, L55 (1984).

[9] J. Karlssonn and A. Roos, Thin Solid Films **392**, 345 (2001).

[10] www.greatvalueconservatories.com.

[11] www.pilkington.com.

[12] www.researchandmarkets.com/reports/c88669.

[13] C. Granqvist, *Handbook of Inorganic Electrochromic Materials* (Elsevier, Amsterdam, 2004).

[14] C. Granqvist, A. Azens, P. Heszler, L. Kish, and L. Österlund, Sol. Energ. Mat. and Sol. C. **91**, 355 (2007).

[15] C. G. Granqvist, Sol. Energ. Mat. Sol. C. **60**, 201 (2000).

[16] S. Deb, S. Lee, C. Tracy, J. Pitts, B. Gregg, and H. Branz, Electrochim. Acta **46**, 2125 (2001).

[17] www.docstoc.com/docs/56265192/Infrared-Reflective-Wall-Paint—Patent-7157112.

[18] www.electronics.frost.com.

[19] H. Ohsaki and Y. Kokubu, Thin Solid Films **351**, 1 (1999).

[20] H. L. Hartnagel, A. K. J. Dawar, and C. Jagadish, *Semiconducting Transparent Thin Films* (Institute of Physics Publishing, Bristol, 1995).

[21] D. S. Ginley and C. Bright, MRS Bull. **25**, 15 (2000).

[22] T. Kamiya and H. Hosono, Int. J. Appl. Ceram. Tec. **2**, 285 (2005).

[23] www.fokus-technologies.de.

[24] H. Hara, T. Shiro, and T. Yatabe, Jpn. J. Appl. Phys. **43**, 745 (2004).

[25] H. Nakazawa, Y. Ito, E. Matsumoto, K. Adachi, N. Aoki, and Y. Ochiai, J. Appl. Phys. **100**, 093706 (2006).

[26] J. R. Bellingham, W. A. Phillips, and C. J. Adkins, J. Phys.: Condens. Matter **2**, 6207 (1990).

[27] M. P. Taylor *et al.*, Adv. Funct. Mater. **18**, 3169 (2008).

[28] T. Minami, T. Kakumu, Y. Takeda, and S. Takata, Thin Solid Films **290-291**, 1 (1996).

[29] D. C. Paine, B. Yaglioglu, Z. Beiley, and S. Lee, Thin Solid Films **516**, 5894 (2008), 5th International Symposium on Transparent Oxide Thin Films for Electronics and Optics.

[30] K. Nomura, H. Ohta, A. Takagi, T. Kamiya, M. Hirano, and H. Hosono, Nature **432**, 488 (2004).

[31] D. C. Look, B. Claflin, Y. I. Alivov, and S. J. Park, Phys. Status Solidi A **201**, 2203 (2004).

[32] D. D. Edwards, T. O. Mason, F. Goutenoire, and K. R. Poeppelmeier, Appl. Phys. Lett. **70**, 1706 (1997).

[33] T. J. Coutts, D. L. Young, X. Li, W. P. Mulligan, and X. Wu, J. Vac. Sci. Technol. A **18**, 2646 (2000).

[34] D. Segev and S.-H. Wei, Phys. Rev. B **71**, 125129 (2005).

[35] H. Hosono, N. Kikuchi, N. Ueda, H. Kawazoe, and K. Shimidzu, Appl. Phys. Lett. **67**, 2663 (1995).

[36] X. Z. Wu, Thin Solid Films **515**, 6000 (2007).

[37] B. Rech, T. Repmann, M. van den Donker, M. Berginski, T. Kilper, J. HÃijpkes, S. Calnan, H. Stiebig, and S. Wieder, Thin Solid Films **511-512**, 548 (2006).

[38] Y. Hamakawa, *Thin-Film Solar Cells: Next Generation Photovoltaics and Its Applications* (Springer-Verlag, Berlin, 2004).

[39] K. L. Chopra, P. D. Paulson, and V. Dutta, Progr. Photovolt: Res. Appl. **12**, 6992 (2004).

[40] www.firstsolar.com/en/modules.php.

[41] C. J. Brabec, N. S. Sariciftci, and J. C. Hummelen, Adv. Funct. Mat. **11**, 15 (2001).

[42] P. Peumans, A. Yakimov, and S. R. Forrest, J. Appl. Phys. **93**, 3693 (2003).

[43] H. Kim, C. M. Gilmore, A. Piqué, J. S. Horwitz, H. Mattoussi, H. Murata, Z. H. Kafafi, and D. B. Chrisey, J. Appl. Phys. **86**, 6451 (1999).

[44] L. S. Hung and C. H. Chen, Mater. Sci. Eng. R **39**, 143 (2002).

[45] J. H. W. Göpel and J. Zemel, *Sensors: A Comprehensive Survey, Chemical and Biochemical Sensors vol. 2* (VCH, Weinheim, 1991).

[46] T. Seiyama, K. Fueki, J. Shiokawa, and S. Suzuki, *Chemical sensors* (Kodansha, Tokyo and Elsevier, Amsterdam, 1983).

[47] M. Batzill and U. Diebold, Physical Chemistry Chemical Physics **9**, 2307 (2007).

[48] N. Taguchi, Patent number: 3644795, Filing date: 17 Nov 1970 (Issue date: Feb 1972).

[49] K. Ihokura and J. Watson, *The Stannic Oxide Gas Sensor - Principles and Applications* (CRC Press, Boca Raton, FL, 1994).

[50] K. D. Schierbaum, H. D. Wiemhöfer, and W. Göpel, Solid State Ionics **28-30**, 1631 (1988).

[51] M. Batzill and U. Diebold, Prog. Surf. Sci. **79**, 47 (2005).

[52] T. Takada, K. Suzuki, and M. Nakane, Sensors and Actuators B: Chemical **13**, 404 (1993).

[53] T. Takada, H. Tanjou, T. Saito, and K. Harada, Sensors and Actuators B: Chemical **25**, 548 (1995), proceedings of the Fifth International Meeting on Chemical Sensors.

[54] A. Gurlo, N. Barsan, M. Ivanovskaya, U. Weimar, and W. Göpel, Sensors and Actuators B: Chemical **47**, 92 (1998).

[55] S.-R. Kim, H.-K. Hong, C. H. Kwon, D. H. Yun, K. Lee, and Y. K. Sung, Sensors and Actuators B: Chemical **66**, 59 (2000).

[56] A. Galdikas, Z. Martunas, and A. Setkus, Sensors and Actuators B: Chemical **7**, 633 (1992).

[57] H. Yamaura, T. Jinkawa, J. Tamaki, K. Moriya, N. Miura, and N. Yamazoe, Sensors and Actuators B: Chemical **36**, 325 (1996), proceedings of the Sixth International Meeting on Chemical Sensors.

[58] I. H. Campbell, D. L. Smith, and J. P. Ferraris, Appl. Phys. Lett. **66**, 3030 (1995).

[59] Y. Zhou, F. Zhang, K. Tvingstedt, S. Barrau, F. Li, W. Tian, and O. Inganas, Appl. Phys. Lett. **92**, 233308 (2008).

[60] S. De and J. N. Coleman, ACS Nano **4**, 2713 (2010).

[61] N. W. Ashcroft and N. D. Mermin, W.B Saunders (1976).

[62] S. Nakamura, M. Senoh, and T. Mukai, Appl. Phys. Lett. **62**, 2390 (1993).

[63] H. Amano, M. Kito, K. Hiramatsu, and I. Akasaki, Jpn. J. Appl. Phys. **28**, L2112 (1989).

[64] S. Fischer, C. Wetzel, E. E. Haller, and B. K. Meyer, Appl. Phys. Lett. **67**, 1298 (1995).

[65] D. R. Lide, *Handbook of Chemistry and Physics* (CRC Press, Boca Raton, London, New York, Washington D.C., 2005).

[66] E. H. Kisi and M. M. Elcombe, Acta Cryst. C **45**, 1867 (1989).

[67] S. Geller, J. Chem. Phys. **33**, 676 (1960).

[68] M. Marezio, Acta Cryst. **20**, 723 (1966).

[69] T. Yamanaka, R. Kurashima, and J. Mimaki, Z. Kristallogr. **215**, 424 (2000).

[70] J. E. Jaffe, R. Pandey, and A. B. Kunz, Phys. Rev. B **43**, 14030 (1991).

[71] A. F. Holleman, E. Wiberg, and N. Wiberg, *Lehrbuch der Anorganischen Chemie* (Walter de Gruyter & Co., Berlin, New York, 1995).

[72] A. Seko, A. Togo, F. Oba, and I. Tanaka, Phys. Rev. Lett. **100**, 045702 (2008).

[73] W. Chen *et al.*, J. Phys. Chem. C **113**, 1320 (2009).

[74] C. W. W. Hoffman, R. C. Ropp, and R. W. Mooney, J. Am. Chem. Soc. **81**, 3830 (1959).

[75] Y. R. Ryu, T. S. Lee, J. H. Leem, and H. W. White, Appl. Phys. Lett. **83**, 4032 (2003).

[76] T. Yamamoto and H. Katayama-Yoshida, Jpn. J. Appl. Phys. .

[77] T. Yamamoto and H. Katayama-Yoshida, Physica B **320-303**, 155 (2001).

[78] T. Yamamoto and H. Katayama-Yoshida, J. Cryst. Growth **214/215**, 552 (2000).

[79] L. G. Wang and A. Zunger, Phys. Rev. Lett. **90**, 256401 (2003).

[80] G. Qin, D. Li, Z. Feng, and S. Liu, Thin Solid Films **517**, 3345 (2009).

[81] J. Stankiewicz, F. Villuendas, and R. Alcalá, Appl. Phys. Lett. **96**, 192108 (2010).

[82] A. Klein, private communication .

[83] S. Lany, J. Osorio-Guillén, and A. Zunger, Phys. Rev. B **75**, 241203 (2007).

[84] S. Wei and A. Zunger, Phys. Rev. B **37**, 8958 (1988).

[85] A. Klein, A. Körber, C. Wachau, F. Säuberlich, Y. Gassenbauer, S. P. Harvey, and T. O. Mason, Thin Solid Films **518**, 1197 (2009).

[86] J. Ghijsen, L. H. Tjeng, J. van Elp, H. Eskes, J. Westerink, G. A. Sawatzky, and M. T. Czyzyk, Phys. Rev. B **38**, 11322 (1988).

[87] K. Matsuzaki, K. Nomura, H. Yanagi, T. Kamiya, M. Hirano, and H. Hosono, Appl. Phys. Lett. **93**, 202107 (2008).

[88] B. Höffling, A. Schleife, F. Fuchs, C. Rödl, and F. Bechstedt, Appl. Phys. Lett. **97**, 032116 (2010).

[89] P. D. C. King, T. D. Veal, A. Schleife, J. Zúñiga-Pérez, B. Martel, P. H. Jefferson, F. Fuchs, V. Muñoz-Sanjosé, F. Bechstedt, and C. F. McConville, Phys. Rev. B **79**, 205205 (2009).

[90] V. Srikant and D. R. Clarke, J. Appl. Phys. **83**, 5447 (1998).

[91] H. H. Tippins, Phys. Rev. **140**, A316 (1965).

[92] N. M. Ravindra, R. A. Weeks, and D. L. Kinser, Phys. Rev. B **36**, 6132 (1987).

[93] P. H. Jefferson, S. A. Hatfield, T. D. Veal, P. D. C. King, C. F. McConville, J. Zúñiga Pérez, and V. Muñoz Sanjosé, Appl. Phys. Lett. **92**, 022101 (2008).

[94] R. L. Weiher and R. P. Ley, J. Appl. Phys. **37**, 299 (1966).

[95] P. D. C. King *et al.*, Phys. Rev. B **79**, 205211 (2009).

[96] A. Bourlange, D. J. Payne, R. G. Egdell, J. S. Foord, P. P. Edwards, M. O. Jones, A. Schertel, P. J. Dobson, and J. L. Hutchison, Appl. Phys. Lett. **92**, 092117 (2008).

[97] F. Fuchs and F. Bechstedt, Phys. Rev. B **77**, 155107 (2008).

[98] E. Kohnke, J. Phys. Chem. of Solids **23**, 1557 (1962).

[99] A. Walsh *et al.*, Phys. Rev. Lett. **100**, 167402 (2008).

[100] J. E. Medvedeva, Euro Phys. Lett. **78**, 57004 (2007).

[101] B. Stjerna, E. Olsson, and C. G. Granqvist, J. Appl. Phys. **76**, 3797 (1994).

[102] H. L. Ma, D. H. Zhang, S. Z. Win, S. Y. Li, and Y. P. Chen, Sol. Energ. Mat. and Sol. C. **40**, 371 (1996).

[103] B. Slater, C. R. A. Catlow, D. H. Gay, D. E. Williams, and V. Dusastre, J. Phys. Chem. B **103**, 10644 (2001).

[104] D. Dobler, S. Oswald, J. Werner, W. Arabczyk, G. Behr, and K. Wetzig, Chem. Phys. **286**, 375 (2003).

[105] O. Warschkow, D. E. Ellis, G. B. González, and T. O. Mason, J. Am. Ceram. Soc. **86**, 1707 (2003).

[106] K. Ellmer, J. Phys. D **33**, R17 (2000).

[107] J. E. Medvedeva and C. L. Hettiarachchi, Phys. Rev. B **81**, 125116 (2010).

[108] P. Agoston, unpublished results .

[109] G. Frank and G. Köstlin, Appl Phys. A **27**, 197 (1982).

[110] G. B. González, T. O. Mason, J. P. Quintana, O. Warschkow, D. E. Ellis, J.-H. Hwang, J. P. Hodges, and J. D. Jorgensen, J. Appl. Phys. **96**, 3912 (2004).

[111] G. B. González, J. B. Cohen, J.-H. Hwang, T. O. Mason, J. P. Hodges, and J. D. Jorgensen, J. Appl. Phys. **89**, 2550 (2001).

[112] O. Warschkow, D. E. Ellis, G. B. González, and T. O. Mason, J. Am. Ceram. Soc. **86**, 1700 (2003).

[113] J. A. Marley and R. C. Dockerty, Phys. Rev. **140**, A304 (1965).

[114] S. Samson and C. G. Fonstad, J. Appl. Phys. **44**, 4618 (1973).

[115] J. Maier and W. Göpel, J. of Solid State Chem. **72**, 293 (1988).

[116] W. Göpel, K. Schierbaum, H. Wiemhöfer, and J. Maier, Solid State Ionics **32-33**, 440 (1989).

[117] C. G. Fonstad, J. Appl. Phys. **42**, 2911 (1971).

[118] G. Rupprecht, Z. Phys. **139**, 504 (1954).

[119] R. L. Weiher, J. Appl. Phys. **33**, 2834 (1962).

[120] J. H. W. de Wit, J. Solid State Chem. **13**, 192 (1975).

[121] J. H. W. de Wit, J. Solid State Chem. **20**, 143 (1977).

[122] J. H. W. de Wit, J. Solid State Chem. **38**, 819 (1977).

[123] J. Mizusaki, H. Koinuma, J. Shimoyama, M. Kawasaki, and K. Fueki, J. Solid State Chem. **88**, 443 (1990).

[124] Y. Li-Zi, S. Zhi-Tong, and W. Chan-Zheng, Solid State Ionics **50**, 203 (1992).

[125] S. Lany and A. Zunger, Phys. Rev. B **72**, 035215 (2005).

[126] C. Kilic and A. Zunger, Phys. Rev. Lett. **88**, 095501 (2002).

[127] K. G. Godinho, A. Walsh, and G. W. Watson, J. Phys. Chem. C **113**, 439 (2009).

[128] I. Tanaka, M. Tatsumi, M. Nkano, and H. Adachi, J. Am. Ceram. Soc. **85**, 68 (2002).

[129] I. Tanaka, F. Oba, K. Tatsumi, M. Kunisu, M. Nkano, and H. Adachi, Mater. Trans. **7**, 1426 (2002).

[130] T. Tomita, K. Yamashita, Y. Hayafuji, and H. Adachi, Appl. Phys. Lett. **87**, 051911 (2005).

[131] S. Lany and A. Zunger, Phys. Rev. Lett. **98**, 045501 (2007).

[132] A. K. Singh, A. Janotti, M. Scheffler, and C. G. V. de Walle, Phys. Rev. Lett. **101**, 055502 (2008).

[133] P. D. Nola, F. Morazzoni, R. Scotti, and D. Narducci, J. Chem. Soc. Faraday T. **89**, 3711 (1993).

[134] S. Limpijumnong, P. Reunchan, A. Janotti, and C. G. Van de Walle, Phys. Rev. B **80**, 193202 (2009).

[135] E. Kane, J. Phys. Chem. Solids **1**, 82 (1956).

[136] C. G. Van de Walle and J. Neugebauer, Nature **423**, 626 (2003).

[137] K. Ellmer, A. Klein, and B. Rech, *Transparent conductive zinc oxide: basics and applications in thin film solar cells* (Springer, Berlin, 2008).

[138] C. Körber, P. Ágoston, and A. Klein, Sensors and Actuators B: Chemical **139**, 665 (2009).

[139] I. Hamberg, C. G. Granqvist, K. F. Berggren, B. E. Sernelius, and L. Engström, Phys. Rev. B **30**, 3240 (1984).

[140] A. J. Leenheer, J. D. Perkins, M. F. A. M. van Hest, J. J. Berry, R. P. O'Hayre, and D. S. Ginley, Phys. Rev. B **77**, 115215 (2008).

[141] www.webelements.com/periodicity/abundance_crust.

[142] K. Salazar and S. M. Kimball, *Mineral Commodity Summaries* (U.S. Department of the Interior and U.S. Geological Survey, Washington, 2009).

[143] www.metalprices.com.

[144] www.aimsolder.com.

[145] U. Schwarz-Schampera and P. M. Herzig, *Indium* (Springer-Verlag, Berlin Heidelberg New York, 2002).

[146] P. Erhart, K. Albe, and A. Klein, Phys. Rev. B **73**, 205203 (2006).

[147] K. Capelle, arXiv **condmat**, 0211443 (2005).

[148] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).

[149] R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).

[150] P. Hohenberg and W. Kohn, Phys. Rev. **136**, 864 (1964).

[151] W. Kohn and L. J. Sham, Appl Phys. A **140**, 1133 (1965).

[152] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[153] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).

[154] G. Kresse and J. Furthmüller, Comp. Mater. Sci. **6**, 15 (1996).

[155] X. Gonze *et al.*, Comp. Mater. Sci. **25**, 478 (2002).

[156] C. Fiolhais, F. Nogueira, and M. Marques, *A Primer in Density Functional Theory* (Springer-Verlag, Berlin Heidelberg New York, 2003).

[157] F. Herman and J. P. Van Dyke, Phys. Rev. Lett **22**, 807 (1969).

[158] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996), erratum: *ibid.* **78**, 1396 (1997).

[159] A. D. Becke, Phys. Rev. A **38**, 13244 (1988).

[160] J. P. Perdew and Y. Wang, Phys. Rev. B **45**, 13244 (1991).

[161] R. M. Martin, *Electronic Structure Basic Theory and Methods* (Cambridge University Press, Cambridge, 2004).

[162] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).

[163] J. Paier, M. Marsman, and G. Kresse, J. Chem. Phys. **127**, 024103 (2007).

[164] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **118**, 8207 (2003).

[165] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **124**, 219906 (2006).

[166] J. Paier, M. Marsman, K. Hummer, G. Kresse, I. C. Gerber, and J. G. Ángyán, J. Chem. Phys. **124**, 154709 (2006).

[167] A. I. Liechtenstein, V. I. Anisimov, and J. Zaanen, Phys. Rev. B **52**, R5467 (1995).

[168] V. I. Anisimov, J. Zaanen, and O. K. Andersen, Phys. Rev. B **44**, 943 (1991).

[169] J. Hubbard, P. R. Soc. London **276**, 238 (1963).

[170] J. M. Soler and A. E., J. Phys. Condens. Matter **14**, 2745 (2002).

[171] K. Ohno, K. Esfarjani, and Y. Kawazoe, Cambridge University Press (2003).

[172] D. R. Hamann, M. Schlüter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1997).

[173] X. Gonze and F. Finocchi, available at http://www.abinit.org (unpublished).

[174] D. Vanderbilt, Phys. Rev. B **41**, 7892 (1990).

[175] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).

[176] D. J. Singh and L. Nordström, *Planewaves, pseudopotentials, and the LAPW method* (Birkhäuser, New York, 2006).

[177] P. Atkins and J. de Paula, *Atkins' Physical Chemistry*, 9th edition. ed. (Oxford University Press, Oxford, 2009).

[178] W. Zhang, J. R. Smith, and X. Wang, Phys. Rev. B **70**, 024103 (2004).

[179] D. C. Wallace, *Thermodynamics of Crystals* (Dover, Mineola, New York, 1998).

[180] D. R. Stull and H. Prohet, *JANAF thermochemical tables 2nd ed.* (U.S. National Bureau of Standards, Washington D.C., 1971).

[181] J. F. Janak, Phys. Rev. B **18**, 7165 (1978).

[182] M. Levy, J. P. Perdew, and V. Sahni, Phys. Rev. A **30**, 2745 (1984).

[183] L. Kleinman, Phys. Rev. B **56**, 16029 (1997).

[184] J. P. Perdew and M. Levy, Phys. Rev. B **56**, 16021 (1997).

[185] L. Kleinman, Phys. Rev. B **56**, 12042 (1997).

[186] T. Koopmans, Physica **1**, 104 (1934).

[187] C. Almbladh and U. von Barth, Phys. Rev. B **31**, 3231 (1985).

[188] L. J. Sham and M. Schlüter, Phys. Rev. B **32**, 3883 (1985).

[189] L. J. Sham and M. Schlüter, Phys. Rev. Lett. **51**, 1888 (1983).

[190] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, Phys. Rev. Lett. **49**, 1691 (1982).

[191] S. Kümmel and L. Kronik, Rev. Mod. Phys. **80**, 3 (2008).

[192] C. Persson, Y.-J. Zhao, S. Lany, and A. Zunger, Phys. Rev. B **72**, 035211 (2005).

[193] G. Makov and M. C. Payne, Phys. Rev. B **51**, 4014 (1995).

[194] N. D. M. Hine, K. Frensch, W. M. C. Foulkes, and M. W. Finnis, Phys. Rev. B **79**, 024112 (2009).

[195] J. Lento, J.-L. Mozos, and R. M. Nieminen, J. Phys.: Condens. Matter **14**, 2637 (2002).

[196] S. Lany and A. Zunger, Phys. Rev. B **78**, 235104 (2008).

[197] J. Shim, E.-K. Lee, Y. J. Lee, and R. M. Nieminen, Phys. Rev. B **71**, 035206 (2005).

[198] C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, Phys. Rev. Lett. **102**, 016402 (2009).

[199] C. W. M. Castleton, A. Höglund, and S. Mirbt, Phys. Rev. B **73**, 035215 (2006).

[200] F. Oba, A. Togo, I. Tanaka, J. Paier, and G. Kresse, Physical Review B (Condensed Matter and Materials Physics) **77**, 245202 (2008).

[201] H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).

[202] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, Rev. Mod. Phys. **73**, 515 (2001).

[203] P. Giannozzi, S. de Gironcoli, P. Pavone, and S. Baroni, Phys. Rev. B **43**, 7231 (1991).

[204] X. Gonze, Phys. Rev. B **55**, 10337 (1997).

[205] G. H. Vineyard, J. Phys. Chem. Solids **3**, 121 (1957).

[206] C. P. Flynn, *Point Defects and Diffusion* (Oxford University Press, Oxford, 1972).

[207] A. R. Allnatt and A. B. Lidiard, *Atomic Transport in Solids* (Cambridge University Press, Cambridge, 2003).

[208] K. A. Fichthorn and W. H. Weinberg, J. Chem. Phys. **95**, 1090 (1991).

[209] G. Henkelman, G. Jóhannesson, and H. Jónsson, in *Methods for finding saddlepoints and minimum energy paths* in *Progress on theoretical chemistry and physics*, p. 269, edited by S. D. Schwartz (Kluwer Academic, Dordrecht, 2000).

[210] G. Henkelman and H. Jónsson, J. Chem. Phys. **113**, 9978 (2000).

[211] G. Henkelman, B. P. Uberuaga, and H. Jónsson, J. Chem. Phys. **113**, 9901 (2000).

[212] K. Reuter and M. Scheffler, Phys. Rev. B **65**, 035406 (2001).

[213] P. W. Tasker, J. Phys. C Solid State **12**, 4977 (1979).

[214] J. Goniakowski, F. Finocchi, and C. Noguera, Rep. Prog. Phys. **71**, 016501 (2008).

[215] C. Noguera, J. Phys.: Condens. Matter **12**, R367 (2000).

[216] R. Nosker, P. Mark, and J. Levine, Surf. Sci. **19**, 291 (1970).

[217] R. M. Martin, Phys. Rev. B **6**, 4546 (1972), *ibid.* **20**, 818 (1979), erratum.

[218] W. A. Harrison, Surf. Sci. **55**, 1 (1976).

[219] W. A. Harrison, E. A. Kraut, J. R. Waldrop, and R. W. Grant, Phys. Rev. B **18**, 4402 (1978).

[220] S. L. Dudarev, G. A. Botton, S. Y. Savrasov, C. J. Humphreys, and A. P. Sutton, Phys. Rev. B **57**, 1505 (1998).

[221] P. Erhart, A. Klein, R. G. Egdell, and K. Albe, Phys. Rev. B **75**, 153205 (2007).

[222] L. F. J. Piper, A. DeMasi, S. W. Cho, K. E. Smith, F. Fuchs, F. Bechstedt, C. Körber, A. Klein, D. J. Payne, and R. G. Egdell, Appl. Phys. Lett. **94**, 022105 (2009).

[223] T. L. Barr and Y. L. Liu, J. Phys. Chem. Solids **50**, 657 (1989).

[224] Y. Mishin, M. R. Sorensen, and A. F. Voter, Philos. Mag. A **81**, 2591 (2001).

[225] J. Fernández, A. Monti, and R. Pasianot, physica status solidi (b) **219**, 245 (2000).

[226] P. W. M. Jacobs, M. A. H. Nerenberg, J. Govindarajan, and T. M. Haridasan, J. Phys. C Solid State **15**, 4245 (1982).

[227] M. J. Gillan and P. W. M. Jacobs, Phys. Rev. B **28**, 759 (1983).

[228] J. H. Harding, Phys. Rev. B **32**, 6861 (1985).

[229] R. D. Hatcher, R. Zeller, and P. H. Dederichs, Phys. Rev. B **19**, 5083 (1979).

[230] E. Rauls and T. Frauenheim, Phys. Rev. B **69**, 155213 (2004).

[231] E. Rauls, T. Frauenheim, A. Gali, and P. Deák, Phys. Rev. B **68**, 155208 (2003).

[232] G. B. Bachelet and G. de Lorenzi, Phys. Scripta **T19A**, 311 (1987).

[233] O. Warschkow, L. Miljacic, D. E. Ellis, G. B. González, and T. O. Mason, J. Amer. Ceramic Soc. **89**, 616 (2006).

[234] O. N. Mryasov and A. J. Freeman, Phys. Rev. B **64**, 233111 (2001).

[235] P. Erhart, A. Klein, and K. Albe, Phys. Rev. B **72**, 085213 (2005).

[236] D. Alfè, (1998), program available at: http://chianti.geol.ucl.ac.uk/∼dario.

[237] *Thermodynamic Properties of Inorganic Materials: Pure Substances. Part 4: Compounds from HgH to ZnTe* (Springer-Verlag, Berlin/Heidelberg, 2001), Vol. 19A4.

[238] K. I. Hagemark, J. Solid State Chem. **16**, 293 (1976).

[239] A. Janotti and C. G. Van de Walle, Phys. Rev. B **76**, 165202 (2007).

[240] V. Anisimov, I. Solovyev, M. Korotin, M. Czyzek, and G. Sawatzky, Phys. Rev. B **46**, 16929 (1993).

[241] A. Liechtenstein and *et. al.*, Phys. Rev. B **52**, R5467 (1995).

[242] C. G. Van de Walle, Phys. Rev. Lett. **85**, 1012 (2000).

[243] Y. Kim and C. H. Park, Phys. Rev. Lett. **102**, 086403 (2009).

[244] J. P. Perdew, M. Ernzerhof, and K. Burke, J. Chem. Phys. **105**, 9982 (1996).

[245] A. Alkauskas, P. Broqvist, and A. Pasquarello, Phys. Rev. Lett. **101**, 046405 (2008).

[246] F. Fuchs, J. Furthmüller, F. Bechstedt, M. Shishkin, and G. Kresse, Phys. Rev. B **76**, 115109 (2007).

[247] S. Lany and A. Zunger, Phys. Rev. Lett. **106**, 069601 (2011).

[248] R. L. Weiher and B. G. Dick, Jr., J. Appl. Phys. **35**, 3511 (1964).

[249] N. Özcan, T. Kortelainen, V. Golovanov, T. T. Rantala, and J. Vaara, Phys. Rev. B **81**, 235202 (2010).

[250] R. Laiho, D. S. Poloskin, Y. P. Stepanov, M. P. Vlasenko, L. S. Vlasenko, and V. S. Zakhvalinskii, J. Appl. Phys. **106**, 013712 (2009).

[251] A. Dixit, C. Sudakar, R. Naik, V. M. Naik, and G. Lawes, Appl. Phys. Lett. **95**, 192105 (2009).

[252] C. A. Pan and T. P. Ma, Appl. Phys. Lett. **37**, 163 (1980).

[253] W. M. Hlaing Oo, S. Tabatabaei, M. D. McCluskey, J. B. Varley, A. Janotti, and C. G. Van de Walle, Phys. Rev. B **82**, 193201 (2010).

[254] H. Iddir, S. Öğüt, P. Zapol, and N. D. Browning, Phys. Rev. B **75**, 073203 (2007).

[255] B. Kamp, R. Merkle, R. Lauck, and J. Maier, J. of Solid State Chem. **178**, 3027 (2005).

[256] J. Gale, J. Chem. Soc **93**, 629 (1997).

[257] J. Gale, Phil. Mag. B **73**, 3 (1996).

[258] P. Ágoston, P. Erhart, A. Klein, and K. Albe, J. Phys. Condens. Matter **21**, 455801 (2009).

[259] P. Ágoston, C. Körber, A. Klein, M. J. Puska, R. M. Nieminen, and K. Albe, J. Appl. Phys. **108**, 053511 (2010).

[260] T. Minami, Semicond. Sci. Technol. **20**, 35 (2005).

[261] K. Ellmer, J. Phys. D **34**, 3097 (2001).

[262] Y. Gassenbauer, R. Schafranek, A. Klein, S. Zafeiratos, M. Hävecker, A. Knop-Gericke, and R. Schlögl, Phys. Rev. B **73**, 245312 (2006).

[263] S. P. Harvey, T. O. Mason, Y. Gassenbauer, R. Schafranek, and A. Klein, J. Phys. D **39**, 3959 (2006).

[264] A. Klein, Appl. Phys. Lett. **77**, 2009 (2000).

[265] J. H. Hwang, D. D. Edwards, D. R. Kammler, and T. O. Mason, Solid State Ionics **129**, 135 (2000).

[266] C. D. Canestraro, M. M. Oliviera, R. Vlasaski, M. V. S. da Silva, D. G. F. David, I. Pepe, A. F. dat Silva, L. S. Roman, and C. Persson, Appl. Surf. Sci. **255**, 1874 (2008).

[267] C. S. Rastomjee, R. G. Egdell, G. C. Geirguadis, M. . J. Lee, and T. J. Tate, J. Mater. Chem. **2**, 511 (1992).

[268] F. J. Berry and B. J. Laundy, J. Chem. Soc.,Dalton Trans 1442 (1981).

[269] D. E. Williams and V. Dusastre, J. Phys. Chem. B **102**, 6732 (1998).

[270] G. P. Wirtz and H. P. Takiar, J. Am. Ceram. Soc. **64**, 748 (1981).

[271] I. Ikuma and T. Murakami, J. Electrochem. Soc. **143**, 2698 (1996).

[272] Y. Ikuma, M. Kamiya, N. Okumura, I. Sakaguchi, H. Haneda, and Y. Sawada, J. Electrochem. Soc. **145**, 2910 (1998).

[273] P. Ágoston and K. Albe, Phys. Chem. Chem. Phys. **11**, 3226 (2009).

[274] P. Ágoston, K. Albe, R. M. Nieminen, and M. J. Puska, Phys. Rev. Lett. **103**, 245501 (2009).

[275] P. Ágoston, K. Albe, R. M. Nieminen, and M. J. Puska, Phys. Rev. Lett. **106**, 069602 (2011).

[276] I. Hamberg and C. G. Granqvist, J. Appl. Phys. **60**, R123 (1986).

[277] C. Xirouchaki, G. Kiriakidis, T. F. Pedersen, and H. Fritzsche, J. Appl. Phys. **79**, 9349 (1996).

[278] J. Baker, M. Muir, and J. Andzelm, J. Chem. Phys. **102**, 2063 (1995).

[279] J. L. Durant, Chem. Phys. Lett. **256**, 595 (1996).

[280] U. Weimar and W. Göpel, Sensors and Actuators B: Chemical **26**, 13 (1995).

[281] W. Göpel and K. D. Schierbaum, Sensors and Actuators B: Chemical **26**, 1 (1995).

[282] J. Jamnik, B. Kamp, R. Merkle, and J. Maier, Solid State Ionics **150**, 157 (2002).

[283] W. Hellmich, C. B. v. Braunmühl, G. Müller, G. Sberveglieri, M. Berti, and C. Perego, Thin Solid Films **263**, 231 (1995).

[284] G. N. Advani, P. Kluge-Weiss, R. Longini, and A. Jordan, Int. J. of Electron. **48**, 403 (1980).

[285] B. Kamp, R. Merkle, and J. Maier, Sensors and Actuators B: Chemical **77**, 534 (2001).

[286] J. P. Perdew, M. Ernzerhof, and K. Burke, J. Chem. Phys. **105**, 9982 (1996).

[287] M. Ernzerhof and G. E. Scuseria, J. Chem. Phys. **110**, 5029 (1999).

[288] C. Adamo and V. Barone, J. Chem. Phys. **110**, 6158 (1999).

[289] P. Erhart and K. Albe, Phys. Rev. B **73**, 115207 (2006).

[290] J. Pohl and K. Albe, J. Appl. Phys. **108**, 023509 (2010).

[291] C. Körber, A. Wachau, P. Ágoston, K. Albe, and A. Klein, Phys. Chem. Chem. Phys. (2011).

[292] T. A. White, M. S. Moreno, and P. A. Midgley, Z. Kristallogr. **225**, 56 (2010).

[293] L. Bursill and B. Hyde, Prog. Solid State Chem. **7**, 177 (1972).

[294] P. Ágoston and K. Albe, Phys. Rev. B **81**, 195205 (2010).

[295] P. Erhart and K. Albe, Appl. Phys. Lett. **88**, 201918 (2006).

[296] E. de Fresart, J. Darville, and J. M. Gilles, Solid State Commun. **37**, 13 (1981).

[297] E. de Fresart, J. Darville, and J. M. Gilles, Appl. Surf. Sci. **11-12**, 637 (1982).

[298] M. Batzill, K. Katsiev, J. M. Burst, U. Diebold, A. M. Chaka, and B. Delley, Phys. Rev. B **72**, 165414 (2005).

[299] F. H. Jones, R. Dixon, J. S. Foord, R. G. Egdell, and J. B. Pethica, Surf.Sci. **376**, 367 (1997).

[300] A. Atrei, E. Zanazzi, U. Bardi, and G. Rovida, Surf.Sci. **475**, L223 (2001).

[301] C. L. Pang, S. A. Haycock, H. Raza, P. J. Møller, and G. Thornton, Phys. Rev. B **62**, R7775 (2000).

[302] D. F. Cox, T. B. Fryberger, and S. Semancik, Phys. Rev. B **38**, 2072 (1988).

[303] D. F. Cox, T. B. Fryberger, and S. Semancik, Surf.Sci. **224**, 121 (1989).

[304] M. Sinner-Hettenbach, M. Göthelid, J. Weissenrieder, H. von Schenk, T. Weiss, N. Barsan, and U. Weimar, Surf. Sci. **477**, 50 (2001).

[305] M. Batzil, K. Kastiev, and U. Diebold, Surf. Sci. **592**, (2003).

[306] J. M. Themlin, R. Sporken, J. Darville, R. Caudano, J. M. Gilles, and R. L. Johnson, Phys. Rev. B **42**, 11914 (1990).

[307] J. Oviedo and M. J. Gillan, Surf.Sci. **463**, 93 (2000).

[308] J. Oviedo and M. J. Gillan, Surf.Sci. **467**, 35 (2000).

[309] A. Beltrán, J. Andrés, E. Longo, and E. R. Leite, Appl. Phys. Lett. **83**, 635 (2003).

[310] W. Bergermayer and I. Tanaka, Appl. Phys. Lett. **84**, 909 (2004).

[311] M. A. Mäkki-Jaskari, T. T. Rantala, and V. Golovanov, Surf.Sci. **577**, 127 (2005).

[312] M. A. Mäki-Jaskari and T. T. Rantala, Phys. Rev. B **65**, 245428 (2002).

[313] T. M. Inerbaev, Y. Kawazoe, and S. Seal, J. Appl. Phys. **107**, 104504 (2010).

[314] J. Oviedo and M. J. Gillan, Surf.Sci. **513**, 26 (2002).

[315] J. Tersoff and D. R. Hamann, Phys. Rev. B **31**, 805 (1985).

[316] A. Gurlo, ChemPhysChem **7**, 2041 (2006).

[317] H. Onishi and Y. Iwasawa, Phys. Rev. Lett. **76**, 791 (1996).

[318] K. T. Park, M. H. Pan, V. Meunier, and E. W. Plummer, Phys. Rev. Lett. **96**, 226105 (2006).

[319] P. Cox, W. Flavell, and R. Egdell, J. Solid State Chem. **68**, 340 (1987).

[320] P. D. C. King, T. D. Veal, D. J. Payne, A. Bourlange, R. G. Egdell, and C. F. McConville, Phys. Rev. Lett. **101**, 116808 (2008).

[321] V. Golovanov, M. A. Mäki-Jaskari, T. T. Rantala, G. Korotcenkov, V. Brinzari, A. Cornet, and J. Morante, Sensors and Actuators B **106**, 563 (2005).

[322] S. P. Harvey, T. O. Mason, C. Körber, Y. Gassenbauer, and A. Klein, Appl. Phys. Lett. **92**, 252106 (2008).

[323] D. J. Milliron, I. G. Hill, C. Shen, A. Kahn, and J. Schwartz, J. Appl. Phys. **87**, 572 (2000).

[324] M. V. Hohman, P. Agoston, A. Wachau, T. Bayer, J. Brötz, K. Albe, and A. Klein, J. Phys.: Condens. Matter **In Press**, (2011).

[325] H. Ohta, M. Orita, M. Hirano, and H. Hosono, J. Appl. Phys. **91**, 3547 (2002).

[326] M. Kamei, H. Enomoto, and I. Yasui, Thin Solid Films **392**, 265 (2001).

[327] Y. Shigesato, S. Takaki, and T. Haranoh, J. Appl. Phys. **71**, 3356 (1992).

[328] A. Bourlange, D. J. Payne, R. G. Palgrave, H. Zhang, J. S. Foord, R. G. Egdell, R. M. J. Jacobs, T. D. Veal, P. D. C. King, and C. F. McConville, J. Appl. Phys. **106**, 013703 (2009).

[329] E. H. Morales and U. Diebold, Appl. Phys. Lett. **95**, 253105 (2009).

[330] Y. Gassenbauer, Ph.D. thesis, Technische Universität Darmstadt, Darmstadt, Germany, 2007.

[331] V. Brinzari, G. Korotcenkov, M. Ivanov, V. Nehasil, V. Matolin, K. Masek, and M. Kamei, Surf.Sci. **601**, 5585 (2007).

[332] E. H. Morales, Y. He, M. Vinnichenko, B. Delley, and U. Diebold, New J. Phys. **10**, 125030 (2008).

[333] E. J. Tarsa, J. H. English, and J. S. Speck, Appl. Phys. Lett. **62**, 2332 (1993).

[334] M. Kamei, T. Yagami, S. Takaki, and Y. Shigesato, Appl. Phys. Lett. **64**, 2712 (1994).

[335] M. Kamei, Y. Shigesato, and S. Takaki, Thin Solid Films **259**, 38 (1995).

[336] O. Bierwagen and J. S. Speck, J. Appl. Phys. **107**, 113519 (2010).

[337] C. Zhou, S. Chen, J. Wu, R. K. Heier, and H. Cheng, J. Phys. Chem. C **112**, 14015 (2008).

[338] Z. R. Xiao, X. F. Fan, L. X. Guan, C. H. A. Huan, J. L. Kuo, and L. Wang, J. Phys.: Condens. Matter **21**, 272202 (2009).

[339] D. Fuks, D. Shapiro, A. Kiv, V. Golovanov, and C. Liu, Int. J. Quantum Chem. (2009).

[340] A. Walsh and C. R. A. Catlow, J. Mater. Chem. (2010).

[341] G. Kresse, O. Dulub, and U. Diebold, Phys. Rev. B **68**, 245409 (2003).

[342] K. Sugiyama, H. Ishii, Y. Ouchi, and K. Seki, J. Appl. Phys. **87**, 295 (2000).

[343] F. Esch, S. Fabris, L. Zhou, T. Montini, C. Africh, P. Fornasiero, G. Comelli, and R. Rosei, Science **309**, 752 (2005).

[344] A. Gurlo, D. Dzivenko, P. Kroll, and R. Riedel, physica status solidi-R **2**, 269 (2008).

[345] D. F. Mullica, G. W. Beall, W. O. Milligan, J. D. Korp, and I. Bernal, J. Inorg. Nucl. Chem. **41**, 277 (1979).

[346] R. Shuttleworth, P. Phys. Soc. A **63**, 444 (1950).

[347] R. Egdell, W. Flavell, and P. Tavener, J. of Solid State Chem. **51**, 345 (1984).

[348] P. Ágoston, C. Körber, M. J. Puska, R. M. Nieminen, A. Klein, and K. Albe, J. Appl. Phys. **108**, 053511 (2010).

[349] K. Hauffe, Angew. Chem. **68**, (1956).

[350] H. Geistlinger, Sensors and Actuators B **17**, 47 (1993).

[351] P. Ágoston and K. Albe, Surf. Sci. **605**, 714 (2011).

[352] Y. Hao, G. Meng, and L. Ye, C. Zhang, Cryst. Growth Desig. **5**, 1617 (2005).

[353] A. Janotti and C. G. Van de Walle, Nat. Mater. **6**, 44 .

[354] J. B. Varley, J. R. Weber, A. Janotti, and C. G. V. de Walle, Appl. Phys. Lett. **97**, 142106 (2010).

[355] P. D. C. King, I. McKenzie, and T. D. Veal, Appl. Phys. Lett. **96**, 062110 (2010).

[356] P. D. C. King *et al.*, Phys. Rev. B **80**, 081201 (2009).

[357] S. Lee, Y.-W. Kim, and H. Chen, Appl. Phys. Lett. **78**, 350 (2001).